CHAPTER-1

MOTHERBOARD & ITS COMPONENT SPECIFIC OBJECTIVES

1.1 CPU – Concept like address lines, data lines, internal registers.
1.2 Modes of operation of CPU – Real mode, IA-32 mode, IA-32 Virtual Real Mode.
1.3 Process Technologies, Dual Independent Bus Architecture, Hyper Threading Technologies & its requirement.
1.4 Processor socket & slots.
1.5 Chipset basic, chipset Architecture, North / South bridge & Hub Architecture.
1.6 Latest chipset for PC
1.7 Overview & features of PCI, PCI –X, PCI express, AGP bus.
1.8 Logical memory organization conventional memory, extended memory, expanded memory.
1.9 Overview & features of SDRAM, DDR, DDR2, DDR3.
1.10 Concept of Cache memory:
1.11 L1 Cache, L2 Cache, L3 Cache, Cache Hit & Cache Miss.
1.13 BIOS – Basics & CMOS Set Up.
1.14 Motherboard Selection Criteria.

CPU – Concept like address lines, data lines, internal registers

Q.What is Bus , Address , data and control Bus

Ans.A collection of wires through which data is transmitted from one part of a computer to another.When used in reference to personal computers, the term *bus* usually refers to *internal bus*. This is a bus that connects all the internal computer components to the CPU and main memory. There's also an expansion bus that enables expansion boards to access the CPU and memory. Also a **bus** is a **common pathway** through which information flows from one component to another. This pathway is used for communication purpose and can be established between two or more computer components.

A bus is capable of being a parallel or serial bus and today all computers utilize two bus types, an **internal bus** or local bus and an **external bus**, also called the **expansion bus**. An internal bus enables a communication between internal components such as a computer video card and memory and an external bus is capable of communicating with external components such as a USB or SCSI device.

A computer or device's bus speed is listed as a MHz, e.g. 100MHz FSB. The throughput of a bus is measured in bits per second or megabytes per second.

Address Bus

It is a group of wires or lines that are used to transfer the addresses of Memory or I/O devices. It is unidirectional. In Intel 8085 microprocessor, Address bus was of 16 bits. This means that Microprocessor 8085 can transfer maximum 16 bit address which means it can address 65,536 different memory locations. This bus is multiplexed with 8 bit data bus. So the most significant bits (MSB) of address goes through Address bus (A7-A0) and LSB goes through multiplexed data bus (AD0-AD7).

Each wire in an address bus carries a single bit of information. This single bit is a single digit in the address. The more wires (digits) used in calculating these addresses, the greater the total number of address locations. The size (or width) of the address bus indicates the maximum amount of RAM a chip can address. The bus contains multiple wires (signal lines) that contain addressing information that describes the memory location of where the data is being sent or where it is being retrieved. Each wire in the bus carries a single bit of information, which means the more wires a bus has the more information it can address. For example, a computer with a 32-bit address bus can address 4GB of memory, and a computer with a 36-bit bus can address 64GB of memory.

64-bit AMD/Intel	
Address Bus	40-bit
Bytes	1,099,511,627,776
KiB	1,073,741,824
MiB	1,048,576
GiB	1024
TiB	1

The data bus and address bus are independent, and chip designers can use whatever size they want for each. Usually, however, chips with larger data buses have larger address buses. The sizes of the buses can provide important information about a chip's relative power, measured in two important ways. The size of the data bus indicates the chip's information-moving capability, and the size of the address bus tells you how much memory the chip can handle.

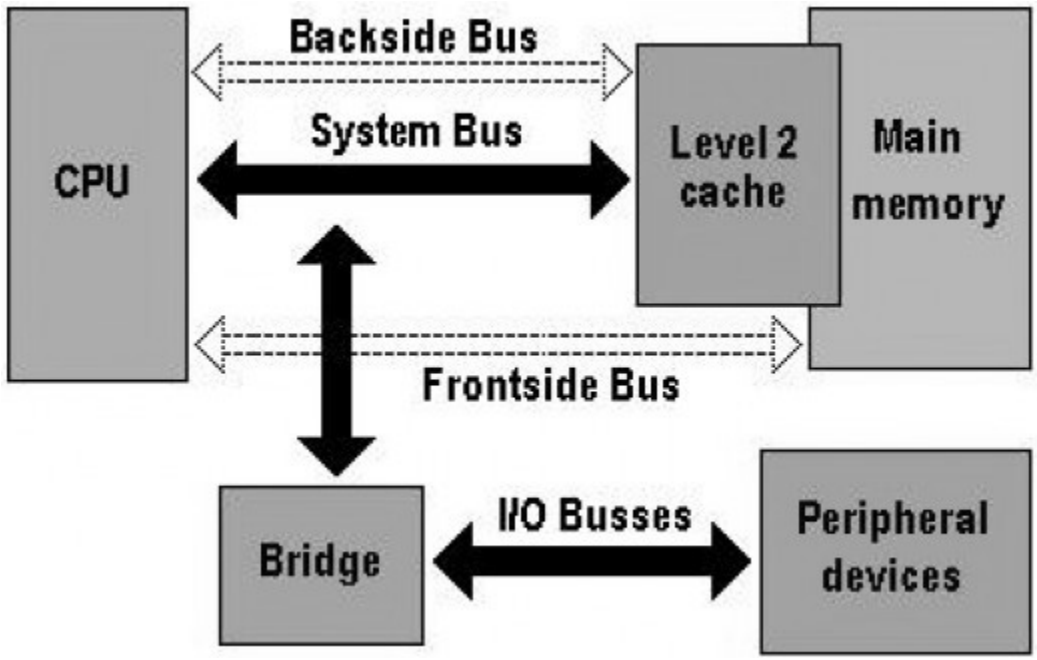
Data Bus:

As name tells that it is used to transfer data within Microprocessor and Memory/Input or Output devices. It is bidirectional as Microprocessor requires to send

or receive data. The data bus also works as address bus when multiplexed with lower order address bus. Data bus is 8 Bits long. The word length of a processor depends on data bus, thats why Intel 8085 is called 8 bit Microprocessor because it have an 8 bit data bus.

To increase the amount of data being sent (called *bandwidth*) by increasing either the cycling time or the number of bits being sent at a time, or both. Over the years, processor data buses have gone from 8 bits wide to 64 bits wide. The more wires you have, the more individual bits you can send in the same interval. All modern processors from the original Pentium and Athlon through the latest Core i7, AMD FX 83xx series, and even the Itanium series have a 64-bit (8-byte)-wide data bus. Therefore, they can transfer 64 bits of data at a time to and from the motherboard chipset or system memory.

Wider the bus more is the speed ie 64-bit-wide buses. Also in newer processors is the use of multiple separate buses for different tasks. Traditional processor design had all the data going through a single bus, whereas newer processors have separate physical buses for data to and from the chipset, memory, and graphics card slot(s).



Control Bus:

Microprocessor uses control bus to process data, that is what to do with the selected memory location. Some control signals are Read, Write and Opcode fetch etc. Various operations are performed by microprocessor with the help of control bus. This is a dedicated bus, because all timing signals are generated according to control signal.

System bus

It is a group of conductors. It is used to transfer information (electrical signal) between two units. It consists of Data Bus, Address Bus and Control Bus.

Functions of Buses

The functions of buses can be summarized as below:

1. **Data sharing** - All types of buses found on a computer must be able to transfer data between the computer peripherals connected to it.
The data is transferred in in either serial or parallel, which allows the exchange of 1, 2, 4 or even 8 bytes of data at a time. (A byte is a group of 8 bits). Buses are classified depending on how many bits they can move at the same time, which means that we have 8-bit, 16-bit, 32-bit or even 64-bit buses.

2. Addressing - A bus has address lines, which match those of the processor. This allows data to be sent to or from specific memory locations.

3. Power - A bus supplies power to various peripherals that are connected to it.

4. Timing - The bus provides a **system clock** signal to synchronize the peripherals attached to it with the rest of the system.

Internal Registers (Internal Data Bus)

The size of the internal registers indicates how much information the processor can operate on at one time and how it moves data around internally within the chip. This is sometimes also referred to as the *internal data bus*. A *register* is a holding cell within the processor; for example, the processor can add numbers in two different registers, storing the result in a third register. The register size determines the size of data on which the processor can operate. The register size also describes the type of software or commands and instructions a chip can run. That is, processors with 32-bit internal registers can run 32-bit instructions that are processing 32-bit chunks of data, but processors with 16-bit registers can't. Processors from the 386 to the Pentium 4 use 32-bit internal registers and can run essentially the same 32-bit OSs and software. The Core 2, Athlon 64, and newer processors have both 32-bit and 64-bit internal registers, which can run existing 32-bit OSs and applications as well as newer 64-bit versions.

A register is a memory location within the CPU itself, designed to be quickly accessed for purposes of fast data retrieval. Processors normally contain a register array, which houses many such registers. These contain instructions, data and other values that may need to be quickly accessed during the execution of a program.

Many different types of registers are common between most microprocessor designs. These are:

Program Counter (PC)

This register is used to hold the memory address of the next instruction that has to executed in a program. This is to ensure the CPU knows at all times where it has reached, that is able to resume following an execution at the correct point, and that the program is executed correctly.

Instruction Register (IR)

This is used to hold the current instruction in the processor while it is being decoded and executed, in order for the speed of the whole execution process to be reduced. This is because the time needed to access the instruction register is much less than continual checking of the memory location itself.

Accumulator (A, or ACC)

The accumulator is used to hold the result of operations performed by the arithmetic and logic unit, as covered in the section on the ALU.

Memory Address Register (MAR)

Used for storage of memory addresses, usually the addresses involved in the instructions held in the instruction register. The control unit then checks this register when needing to know which memory address to check or obtain data from.

Memory Buffer Register (MBR)

When an instruction or data is obtained from the memory or elsewhere, it is first placed in the memory buffer register. The next action to take is then determined and carried out, and the data is moved on to the desired location.

Flag register / status flags

The flag register is specially designed to contain all the appropriate 1-bit status flags, which are changed as a result of operations involving the arithmetic and logic unit. Further information can be found in the section on the ALU.

Index register

A hardware element which holds a number that can be added to (or, in some cases, subtracted from) the address portion of a computer instruction to form an effective address. Also known as **base register**. An index register in a computer's CPU is a processor register used for modifying operand addresses during the run of a program.

Other general purpose registers

These registers have no specific purpose, but are generally used for the quick storage of pieces of data that are required later in the program execution. In the model used here these are assigned the names A and B, with suffixes of L and U indicating the lower and upper sections of the register respectively.

Modes of Operation of CPU

Q.List and Describe Modes of Operation of CPU

Ans. All Intel and Intel-compatible processors from the 386 on up can run in several modes. Processor modes refer to the various operating environments and affect the instructions and capabilities of the chip. The processor mode controls how the processor sees and manages the system memory and the tasks that use it.

The following table summarizes the processor modes and submodes:

Mode	Real	IA-32	IA-32e
Submode	N/A	Protected Virtual real	64-bit compatibility
OS Required	16-bit	32-bit 32-bit	64-bit 64-bit
Software	16-bit	32-bit 16-bit	64-bit 32-bit
Memory Address Size	24-bit	32-bit 24-bit	64-bit 32-bit
Default Operand Size	16-bit	32-bit 16-bit	32-bit 32-bit
Register Width	16-bit	32/16-bit 16-bit	64-bit 32-16-bit
*IA-32e (64-bit extension mode) is also called x64, AMD64, x86-64, or EM64T.			

Real Mode

Real mode is sometimes called 8086 mode because it is based on the 8086 and 8088 processors. The original IBM_PC included an 8088 processor that could execute 16-

bit instructions using 16-bit internal registers and could address only 1 MB of memory using 20 address lines. All original PC software was created to work with this chip and was designed around the 16-bit instruction set and 1 MB memory model. For example, DOS and all DOS software, Windows 1.x through 3.x, and all Windows 1.x through 3.x applications are written using 16-bit instructions. These 16-bit OSs and applications are designed to run on an original 8088 processor.

Later processors such as the 286 could run the same 16-bit instructions as the original 8088, but much faster. In other words, the 286 was fully compatible with the original 8088 and could run all 16-bit software just the same as an 8088, but, of course, that software would run faster. The 16-bit instruction mode of the 8088 and 286 processors has become known as *real mode*. All software running in real mode must use only 16-bit instructions and live within the 20-bit (1 MB) memory architecture it supports. Software of this type is usually single-tasking—that is, only one program can run at a time. No built-in protection exists to keep one program from overwriting another program or even the OS in memory. Therefore, if more than one program is running, one of them could bring the entire system to a crashing halt.

IA-32 (32-Bit) : Protected Mode

Intel 386 was the PC industry's first 32-bit processor. This chip could run an entirely new 32-bit instruction set. To take full advantage of the 32-bit instruction set, a 32-bit OS and a 32-bit application were required. This new 32-bit mode was referred to as *protected mode*, which alludes to the fact that software programs running in that mode are protected from overwriting one another in memory. Such protection makes the system much more crash-proof because an errant program can't easily damage other programs or the OS. In addition, a crashed program can be terminated while the rest of the system continues to run unaffected.

Knowing that new OSs and applications—which take advantage of the 32-bit protected mode—would take some time to develop, Intel wisely built a backward-compatible real mode into the 386. That enabled it to run unmodified 16-bit OSs and applications. It ran them quite well—much more quickly than any previous chip. For most people, that was enough. They did not necessarily want new 32-bit software; they just wanted their existing 16-bit software to run more quickly. Unfortunately, that meant the chip was never running in the 32-bit protected mode, and all the features of that capability were being ignored.

When a 386 or later processor is running DOS (real mode), it acts like a "Turbo 8088," which means the processor has the advantage of speed in running any 16-bit

programs; it otherwise can use only the 16-bit instructions and access memory within the same 1 MB memory map of the original 8088. Therefore, if you have a system with a current 32-bit or 64-bit processor running Windows 3.x or DOS, you are effectively using only the first megabyte of memory, leaving all the other RAM largely unused! New OSs and applications that ran in the 32-bit protected mode of the modern processors were needed.

Note : Windows_XP was the first true 32-bit OS that became a true mainstream product, and that is primarily because Microsoft coerced us in that direction with Windows 9x/Me (which are mixed 16-bit/32-bit systems). Windows 3.x was the last 16-bit OS, which some did not really consider a complete OS because it ran on top of DOS.

IA-32 Virtual Real Mode

The key to the backward compatibility of the Windows 32-bit environment is the third mode in the processor **: virtual real mode**. *Virtual real* is essentially a virtual real mode 16-bit environment that runs inside 32-bit protected mode. When you run a DOS prompt window inside Windows, you have created a virtual real mode session. Because protected mode enables true multitasking, you can actually have several real mode sessions running, each with its own software running on a virtual PC. These can all run simultaneously, even while other 32-bit applications are running.

Note : any program running in a virtual real mode window can access up to only 1MB of memory, which that program will believe is the first and only megabyte of memory in the system. In other words, if you run a DOS application in a virtual real window, it will have a 640 KB limitation on memory usage. That is because there is only 1 MB of total RAM in a 16-bit environment, and the upper 384KB is reserved for system use. The virtual real window fully emulates an 8088 environment, so that aside from speed, the software runs as if it were on an original real mode–only PC. Each virtual machine gets its own 1 MB address space, an image of the real hardware basic input/output system (BIOS) routines, and emulation of all other registers and features found in real mode.

Virtual real mode is used when you use a DOS window to run a DOS or Windows 3.x 16-bit program. When you start a DOS application, Windows creates a virtual DOS machine under which it can run.

Note : All Intel and Intel-compatible (such as AMD and VIA/Cyrix) processors power up in real mode. If you load a 32-bit OS, it automatically switches the processor into 32-bit mode and takes control from there.

It's also important to note that some 16-bit (DOS and Windows 3.x) applications misbehave in a 32-bit environment, which means they do things that even virtual real mode does not support. Diagnostics software is a perfect example of this. Such software does not run properly in a real mode (virtual real) window under Windows. In that case, you can still run your modern system in the original no-frills real mode by booting to a DOS or Windows 9x/Me startup floppy or by using a self-booting CD or DVD that contains the diagnostic software.

Although 16-bit DOS and "standard" DOS applications use real mode, special programs are available that "extend" DOS and allow access to extended memory (over 1 MB). These are sometimes called *DOS extenders* and usually are included as part of any DOS or Windows 3.x software that uses them. The protocol that describes how to make DOS work in protected mode is called *DOS protected mode interface* (DPMI).
Windows 3.x used DPMI to access extended memory for use with Windows 3.x applications. It allowed these programs to use more memory even though they were still 16-bit programs. DOS extenders are especially popular in DOS games because they enable them to access much more of the system memory than the standard 1 MB that most real mode programs can address. These DOS extenders work by switching the processor in and out of real mode. In the case of those that run under Windows, they use the DPMI interface built into Windows, enabling them to share a portion of the system's extended memory.

Another exception in real mode is that the first 64 KB of extended memory is actually accessible to the PC in real mode, despite the fact that it's not supposed to be possible. This is the result of a bug in the original IBM AT with respect to the 21st memory address line, known as *A20* (A0 is the first address line). By manipulating the A20 line, real mode software can gain access to the first 64 KB of extended memory —the first 64 KB of memory past the first megabyte. This area of memory is called the *high memory area* (HMA).

IA-32e 64-Bit Extension Mode (x64, AMD64, x86

64-bit extension mode is an enhancement to the IA-32 architecture originally designed by AMD and later adopted by Intel.

In 2003, AMD introduced the first 64-bit processor for x86-compatible desktop computers—the Athlon 64—followed by its first 64-bit server processor, the Opteron. In 2004, Intel introduced a series of 64-bit-enabled versions of its Pentium 4 desktop processor. The years that followed saw both companies introducing more and more processors with 64-bit capabilities.

Processors with 64-bit extension technology can run in real (8086) mode, IA-32 mode, or IA-32e mode. IA-32 mode enables the processor to run in protected mode and virtual real mode. IA-32e mode allows the processor to run in 64-bit mode and compatibility mode, which means you can run both 64-bit and 32-bit applications simultaneously. IA-32e mode includes two submodes:

- **64-bit mode**—Enables a 64-bit OS to run 64-bit applications
- **Compatibility mode**—Enables a 64-bit OS to run most existing 32-bit software

IA-32e 64-bit mode is enabled by loading a 64-bit OS and is used by 64-bit applications. In the 64-bit submode, the following new features are available:

- 64-bit linear memory addressing
- Physical memory support beyond 4GB (limited by the specific processor)
- Eight new general-purpose registers (GPRs)
- Eight new registers for streaming SIMD extensions (MMX, SSE, SSE2, and SSE3)
- 64-bit-wide GPRs and instruction pointers

IE-32e compatibility mode enables 32-bit and 16-bit applications to run under a 64-bit OS. Unfortunately, legacy 16-bit programs that run in virtual real mode (that is, DOS programs) are not supported and will not run, which is likely to be the biggest problem for many users, especially those that rely on legacy business applications or like to run very old games. Similar to 64-bit mode, compatibility mode is enabled by the OS on an individual code basis, which means 64-bit applications running in 64-bit mode can operate simultaneously with 32-bit applications running in compatibility mode.

What we need to make all this work is a 64-bit OS and, more importantly, 64-bit drivers for all our hardware to work under that OS. Although Microsoft released a 64-bit version of Windows XP, few companies released 64-bit XP drivers. It wasn't until Windows Vista and especially Windows 7 x64 versions were released that 64-bit drivers became plentiful enough that 64-bit hardware support was considered mainstream.

Note : Microsoft uses the term *x64* to refer to processors that support either AMD64 or EM64T because AMD and Intel's extensions to the standard IA32 architecture are practically identical and can be supported with a single version of Windows.
Note: Early versions of EM64T-equipped processors from Intel lacked support for the LAHF and SAHF instructions used in the AMD64 instruction set. However,Pentium 4 and Xeon DP processors using core steppings G1 and higher completely support these

instructions; a BIOS update is also needed. Newer multicore processors with 64-bit support include these instructions as well.

The physical memory limits for Windows XP and later are shown in the table below:

Windows Version	Memory Limit
8 Enterprise/Professional	512 GB
8	128 GB
7 Profession/Ultimate/Enterprise	192 GB
Vista Business/Ultimate/Enterprise	128 GB
Vista/7 Home Premium	16 GB
Vista/7 Home Basic	8 GB
XP Professional	128 GB
XP Home	4 GB

The major difference between 32-bit and 64-bit Windows is memory support—specifically, breaking the 4 GB barrier found in 32-bit Windows systems. 32-bit versions of Windows support up to 4 GB of physical memory, with up to 2 GB of dedicated memory per process. 64-bit versions of Windows support up to 512 GB of physical memory, with up to 4 GB for each 32-bit process and up to 8 TB for each 64-bit process. Support for more memory means applications can preload more data into memory, which the processor can access much more quickly.

64-bit Windows runs 32-bit Windows applications with no problems, but it does not run 16-bit Windows, DOS applications, or any other programs that run in virtual real mode. Drivers are another big problem. 32-bit processes cannot load 64-bit dynamic link libraries (DLLs), and 64-bit processes cannot load 32-bit DLLs. This essentially means that, for all the devices you have connected to your system, you need both 32-bit and 64-bit drivers for them to work. Acquiring 64-bit drivers for older devices or devices that are no longer supported can be difficult or impossible. Before installing a 64-bit version of Windows, be sure to check with the vendors of your internal and add-on hardware for 64-bit drivers.

Although vendors have ramped up their development of 64-bit software and drivers, you should still keep all the memory size, software, and driver issues in mind when considering the transition from 32-bit to 64-bit technology. The transition from

32-bit hardware to mainstream 32-bit computing took 16 years. The first 64-bit PC processor was released in 2003, and 64-bit computing really didn't become mainstream until the release of Windows 7 in late 2009.

Summary

Real Mode

- Like 8086/88 processor
- Use only 16 bit features
- Operates in DOS operation system
- Use only 8086 instruction set
- Uses 16-bit base and offset registers
- Access only 1Mb of physical memory
- All IA-32 processor initialize into real mode
- Concept of segmentation is used.

Protected Mode

- Uses full 32bit feature of the processor
- Process 32 bit instruction
- Can access upto 4Gb of memory
- Uses 32 bit internal registers
- Used by windows , Linux , Os2 operating system
- Concept paging is used

Virtual Real Mode

- Processor runs in protected mode, but simulates real mode: a 20-bit linear address is translated by paging to a 32-bit physical address.
- A processor is switched to virtual mode when running a DOS application under Windows operating system.

Q.Difference between real protected and Virtual Mode

Students have to solve the above question

Process technologies

Q.List and explain different Process technologies

Ans. 1. Dual Independent Bus Architecture

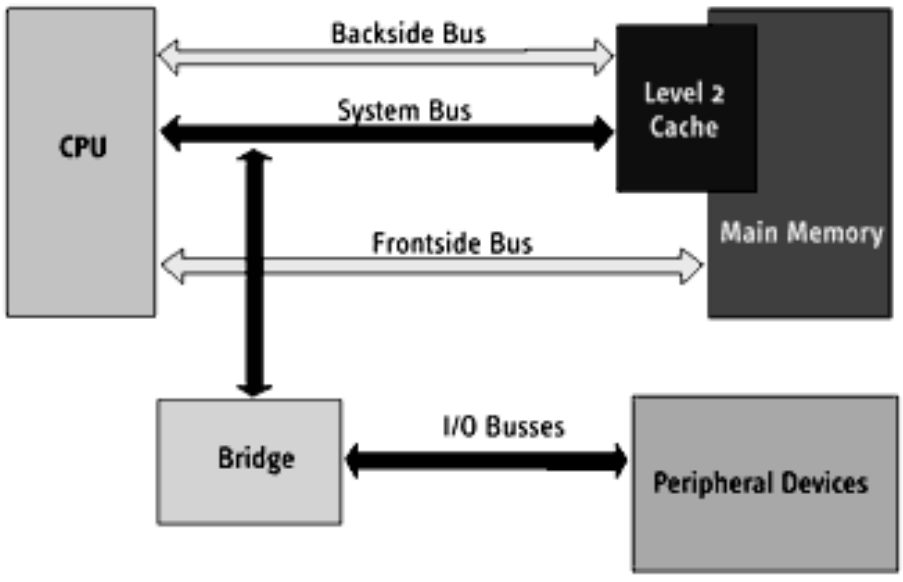
2. Hyper threading

3.MutliCore

Dual Independent Bus – Architecture

The Dual Independent Bus (DIB) architecture was first implemented in the sixth-generation processors from Intel and AMD. DIB was created to improve processor bus bandwidth and performance. Having two (dual) independent data I/O buses enables the processor to access data from either of its buses simultaneously and in parallel, rather than in a singular sequential manner (as in a single-bus system). The main (often called *front-side*) processor bus is the interface between the processor and the motherboard or chipset. The second (back-side) bus in a processor with DIB is used for the L2 cache, enabling it to run at much greater speeds than if it were to share the main processor bus.

Two buses make up the DIB architecture: the L2 cache bus and the main CPU bus, often called *FSB (front side bus)*. The P6 class processors, from the Pentium Pro to the Core 2, as well as Athlon 64 processors can use both buses simultaneously, eliminating a bottleneck there. The dual bus architecture enables the L2 cache of the newer processors to run at full speed inside the processor core on an independent bus, leaving the main CPU bus (FSB) to handle normal data flowing in and out of the chip. The two buses run at different speeds. The front-side bus or main CPU bus is coupled to the speed of the motherboard, whereas the back-side or L2 cache bus is coupled to the speed of the processor core. As the frequency of processors increases, so does the speed of the L2 cache.



DIB also enables the system bus to perform multiple simultaneous transactions (instead of singular sequential transactions), accelerating the flow of information within the system and

boosting performance. Overall, DIB architecture offers up to three times the bandwidth performance over a single-bus architecture processor.

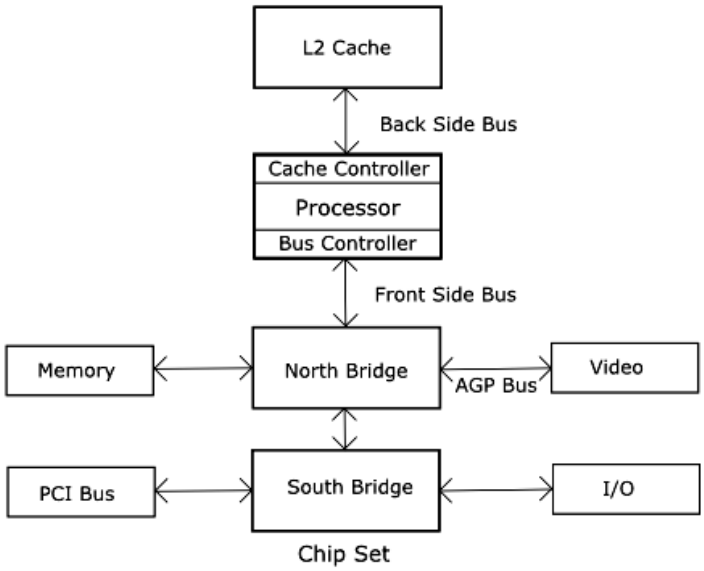


Fig.

Advantages of DIB

1.Faster cache Access
2.Improves Band Width
3.Bot busses are accessed simultaneously hence through put is improved
4.Allow Multiple simultaneous cache request.

Hyper Threading – Intel Proprietary

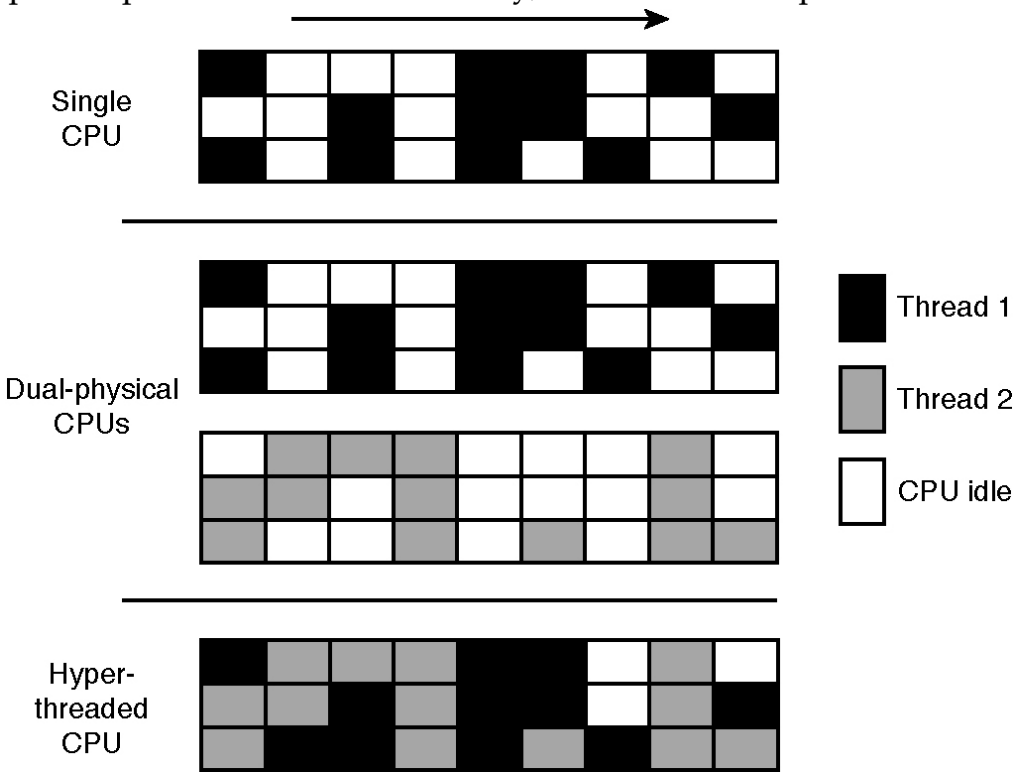
Intel's HT Technology allows a single processor or processor core to handle two independent sets of instructions at the same time. In essence, HT Technology converts a single physical processor core into two virtual processors.

The point of hyper threading is that many times when you are executing code in the processor, there are parts of the processor that is idle. By including an extra set of CPU registers, the processor can act like it has two cores and thus use all parts of the processor in parallel. When the 2 cores both need to use one component of the processor, then one core ends up waiting of course. This is why it can not replace dual-core and such processors.

Hyper-Threading is a technology used by some Intel microprocessor s that allows a single microprocessor to act like two separate processors to the operating system and the application program s that use it. It is a feature of Intel's IA-32 processor architecture.

With Hyper-Threading, a microprocessor's "core" processor can execute two (rather than one) concurrent streams (or thread s) of instructions sent by the operating system. Having two streams of execution units to work on allows more work to be done by the processor during each clock cycle . To the operating system, the Hyper-Threading microprocessor appears to be two separate processors. Because most of today's operating systems (such as Windows and Linux) are capable of dividing their work load among multiple processors (this is called symmetric multiprocessing or SMP), the operating system simply acts as though the Hyper-Threading processor is a pool of two processors.

HT Technology was introduced on Xeon workstation-class processors with a 533 MHz system bus in March 2002. It found its way into standard desktop PC processors starting with the Pentium 4 3.06 GHz processor in November 2002. HT Technology predates multicore processors, so processors that have multiple physical cores, such as the Core 2 and Core i Series, may or may not support this technology depending on the specific processor version. A quad-core processor that supports HT Technology (like the Core i Series) would appear as an 8-core processor to the OS; Intel's Core i7-3970X has six cores and supports up to 12 threads. Internally, an HT-enabled processor has two sets of general-purpose registers, control registers, and other architecture components for each core, but both logical processors share the same cache, execution units, and buses. During operations, each logical processor handles a single thread.



A processor with HT Technology enabled can fill otherwise-idle time with a second process for each core, improving multitasking and performance of multithreading single applications.

Although the sharing of some processor components means that the overall speed of an HT-enabled system isn't as high as a processor with as many physical cores would be, speed increases of 25% or more are possible when multiple applications or multithreaded applications are being run.

To take advantage of HT Technology, you need the following:

- **Processor supporting HT Technology**—This includes many (but not all) Core i Series, Pen-tium 4, Xeon, and Atom processors. Check the specific model processor specifications to be sure.
- **Compatible chipset**—Some older chipsets may not support HT Technology**.**
- **BIOS support to enable/disable HT Technology**—Make sure you enable HT Technology in the BIOS Setup.
- **HT Technology-enabled OS**—Windows XP and later support HT Technology. Linux distributions based on kernel 2.4.18 and higher also support HT Technology. To see if HT Technology is functioning properly, you can check the Device Manager in Windows to see how many processors are recognized. When HT is supported and enabled, the Windows Device Manager shows twice as many processors as there are physical processor cores.

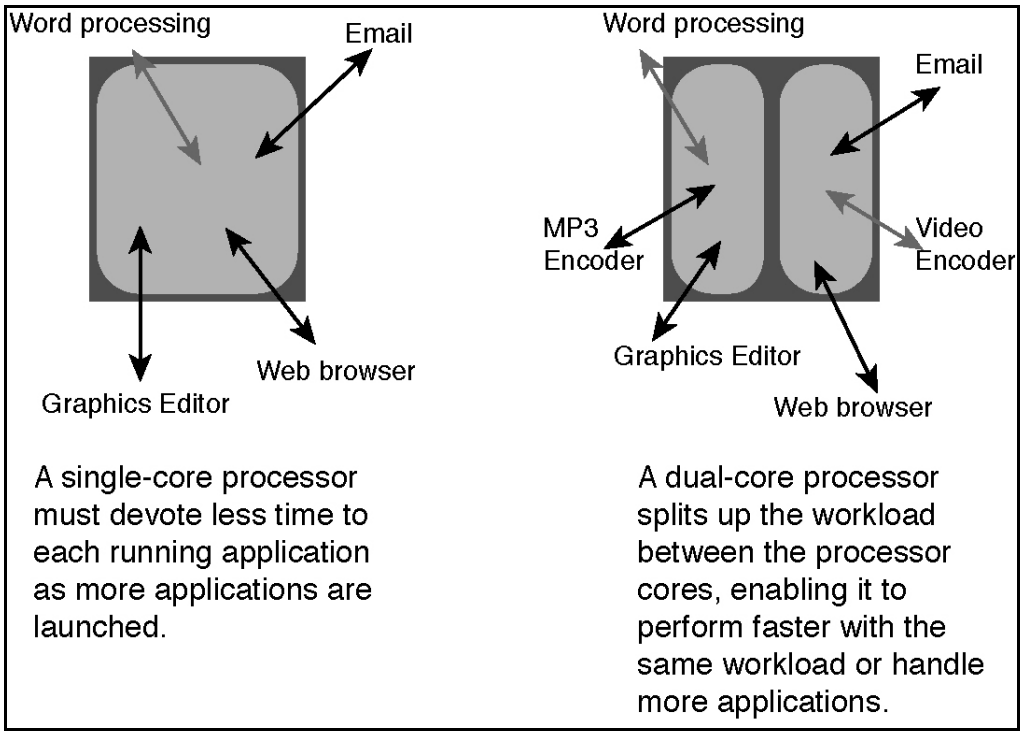
Multicore Technology

HT Technology *simulates* two processors in a single physical core. If multiple logical processors are good, having two or more *physical* processors is a lot better. A multi-core processor, as the name implies, actually contains two or more processor cores in a single processor package. From outward appearances, it still looks like a single processor (and is considered as such for Windows licensing purposes), but inside there can be two, three, four, or even more processor cores. A multi-core processor provides virtually all the advantages of having multiple separate physical processors, all at a much lower cost.

Both AMD and Intel introduced the first dual-core x86-compatible desktop processors in May 2005. AMD's initial entry was the Athlon 64 X2, whereas Intel's first dual-core processors were the Pentium Extreme Edition 840 and the Pentium D. The Extreme Edition 840 was notable for also supporting HT Technology, allowing it to

appear as a quad-core processor to the OS. These processors combined 64-bit instruction capability with dual internal cores—essentially two processors in a single package. These chips were the start of the multicore revolution, which has continued by adding more cores along with additional extensions to the instruction set. Intel introduced the first quad-core processors in November 2006, called the Core 2 Extreme QX and Core 2 Quad. AMD subsequently introduced its first quad-core desktop PC processor in November 2007, called the Phenom.

Note: There has been some confusion about Windows and multi-core or Hyper-Threaded processors. Windows XP and later Home editions support only one physical CPU, whereas Windows Professional, Business, Enterprise, and Ultimate editions support two physical CPUs. Even though the Home editions support only a single physical CPU, if that chip is a multicore processor with HT Technology, all the physical and virtual cores are supported. For example, if you have a system with a quad-core processor supporting HT Technology, Windows Home editions will see it as eight processors, and all of them will be supported. If you had a motherboard with two of these CPUs installed, Windows Home editions would see the eight physical/virtual cores in the first CPU, whereas Professional, Business, Enterprise, and Ultimate editions would see all 16 cores in both CPUs.


Word processing
Email
Web browser
Graphics Editor
A single-core processor must devote less time to each running application as more applications are launched.
Word processing
Email
MP3 Encoder
Video Encoder
Graphics Editor
Web browser
A dual-core processor splits up the workload between the processor cores, enabling it to perform faster with the same workload or handle more applications.

Multi-core processors are designed for users who run multiple programs at the same time or who use multithreaded applications, which pretty much describes all users these days. A multithreaded application can run different parts of the program, known as *threads*, at the same time in the same address space, sharing code and data. A multithreaded program runs faster on a multicore processor or a processor with HT Technology enabled than on a single-core or non-HT processor.

The diagram below illustrates how a single-core processor (left) and a dual-core processor (right) handle multitasking:

It's important to realize that multicore processors don't improve single-task performance much. If you play non-multithreaded games on your PC, it's likely that you would see little advantage in a multi-core or hyperthreaded CPU. Fortunately, more and more software (including games) is designed to be multithreaded to take advantage of multi-core processors. The program is broken into multiple threads, all of which can be divided among the available CPU cores.

Q.State Difference between Hyper threading and Multicore Processor technology

Processor Slot and Sockets

Q.Write short note on Processor Slot and Sockets.
Ans.

CPU Socket

A **CPU socket** or CPU slot is a mechanical component(s) that provides mechanical and electrical connections between a microprocessor and a printed circuit board (PCB). This allows the CPU to be replaced without soldering.

Common sockets have retention clips that apply a constant force, which must be overcome when a device is inserted. For chips with a large number of pins, either zero insertion force (ZIF) sockets or land grid array (LGA) sockets are used instead. These designs apply a compression force once either a handle (for ZIF type) or a surface plate (LGA type) is put into place. This provides superior mechanical retention while avoiding the risk of bending pins when inserting the chip into the socket.

CPU sockets are used in desktop and server computers. As they allow easy swapping of components, they are also used for prototyping new circuits. Laptops typically use surface mount CPUs, which need less space than a socketed part.

Function

A CPU socket is made of plastic, a lever or latch, and metal contacts for each of the pins or lands on the CPU. Many packages are keyed to ensure the proper insertion of the CPU. CPUs with a PGA (pin grid array) package are inserted into the socket and the latch is closed. CPUs with an LGA package are inserted into the socket, the latch plate is flipped into position atop the CPU, and the lever is lowered and locked into place, pressing the CPU's contacts firmly against the socket's lands and ensuring a good connection, as well as increased mechanical stability.

http://www.computerhope.com/jargon/s/socket.htm

http://www.tomshardware.com/reviews/processors-cpu-apu-features-upgrade,3569-15.html

Processor Slot

A **slot** is a computer processor connection designed to make upgrading the processor much easier, where the user would only have to slide a processor into a slot.

The original slot, or **Slot 1** (pictured below), was first released by the Intel Corporation in 1997 as a successor to the Socket 8. Later, AMD released another slot processor known as the **Slot A** in 1999. Both slots look similar but are not compatible. Later, Intel released the slot 2, which was a bigger slot used with the later versions of the Pentium II processors. Today, slot processors are no longer found in new computers and have been replaced by sockets.

A slot **is another name for an expansion slot such as a ISA, PCI, AGP slot, or memory slots.**

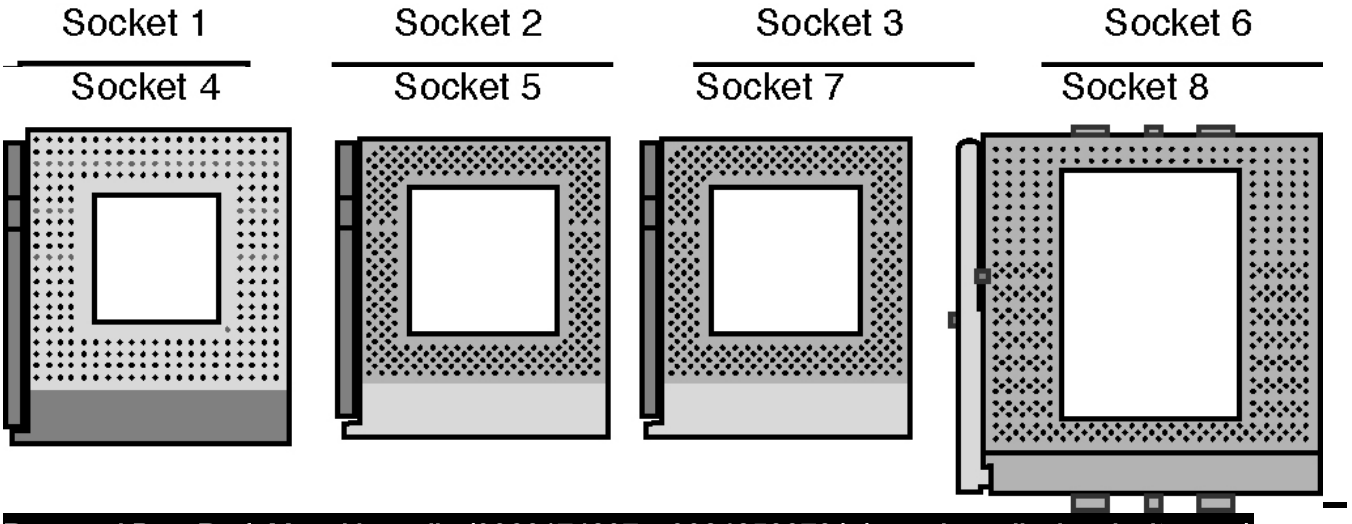
The other form that processors take is a chip soldered on to a card, which then connects to a motherboard by a slot similar to an expansion slot. The picture slows a slot for a **Pentium 3** processor.

Processor Socket and Slot Types

Intel and AMD have created a set of socket and slot designs for their processors. Each socket or slot is designed to support a different range of original and upgrade processors. Table 3.18 shows the designations for the various 486 and newer processor sockets/slots and lists the chips designed to plug into them.

Sockets 1, 2, 3, and 6 are 486 processor sockets and are shown together in Figure so you can see the overall size comparisons and pin arrangements between these sockets. Sockets 4, 5, 7, and 8 are Pentium and Pentium Pro processor sockets and are shown together in Figure so you can see the overall size comparisons and pin arrangements between these sockets. More detailed drawings of each socket are included throughout the remainder of this section with thorough descriptions of the sockets.

486 processor sockets.



Pentium and Pentium Pro processor sockets.

Zero Insertion Force

When the Socket 1 specification was created, manufacturers realized that if users were going to upgrade processors, they had to make the process easier. The socket manufacturers found that 100 lbs. of insertion force is required to install a chip in a standard 169-pin screw Socket 1 motherboard. With this much force involved, you easily could damage either the chip or the socket during removal or reinstallation. Because of this, some motherboard manufacturers began using low insertion force (LIF) sockets, which required only 60 lbs. of insertion force for a 169-pin chip. With the LIF or standard socket, I usually advise removing the motherboard—that way you can support the board from behind when you insert the chip. Pressing down on the motherboard with 60–100 lbs. of force can crack the board if it is not supported properly. A special tool is also required to remove a chip from one of these sockets. As you can imagine, even the low insertion force was relative, and a better solution was needed if the average person was ever going to replace his CPU.

Manufacturers began using ZIF sockets in Socket 1 designs, and all processor sockets from Socket 2 and higher have been of the ZIF design. ZIF is required for all the higher-density sockets because the insertion force would simply be too great otherwise. ZIF sockets almost eliminate the risk involved in installing or removing a processor because no insertion force is necessary to install the chip and no tool is needed to extract one. Most ZIF sockets are handle-actuated: You lift the handle, drop the chip into the socket, and then close the handle. This design makes installing or removing a processor an easy task.

Socket 1

The original OverDrive socket, now officially called Socket 1, is a 169-pin PGA socket. Motherboards that have this socket can support any of the 486SX, DX, and DX2 processors and the DX2/OverDrive versions. This type of socket is found on most 486 systems that originally were designed for OverDrive upgrades. Figure shows the pinout of Socket 1.

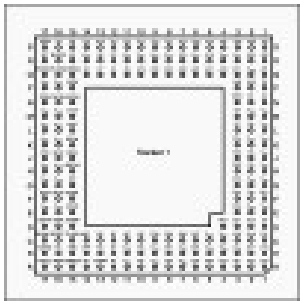


Figure Intel Socket 1 pinout.

The original DX processor draws a maximum 0.9 amps of 5V power in 33MHz form (4.5 watts) and a maximum 1 amp in 50MHz form (5 watts). The DX2 processor, or OverDrive processor, draws a maximum 1.2 amps at 66MHz (6 watts). This minor increase in power requires only a passive heatsink consisting of aluminum fins that are glued to the processor with thermal transfer epoxy. Passive heatsinks don't have any mechanical components like fans. Heatsinks with fans or other devices that use power are called *active* heatsinks. OverDrive processors rated at 40MHz or less do not have heatsinks.

Socket 2

When the DX2 processor was released, Intel was already working on the new Pentium processor. The company wanted to offer a 32-bit, scaled-down version of the Pentium as an upgrade for systems that originally came with a DX2 processor. Rather than just increasing the clock rate, Intel created an allnew chip with enhanced capabilities derived from the Pentium.

The chip, called the Pentium OverDrive processor, plugs into a processor socket with the Socket 2 or Socket 3 design. These sockets hold any 486 SX, DX, or DX2 processor, as well as the Pentium OverDrive. Because this chip is essentially a 32-bit version of the (normally 64-bit) Pentium chip, many have taken to calling it a Pentium-SX. It was available in 25/63MHz and 33/83MHz versions. The first number indicates the base motherboard speed; the second number indicates the actual operating speed of the Pentium OverDrive chip. As you can see, it is a clock-multiplied chip that runs at 2.5 times the motherboard speed. Figure shows the pinout configuration of the official Socket 2 design.

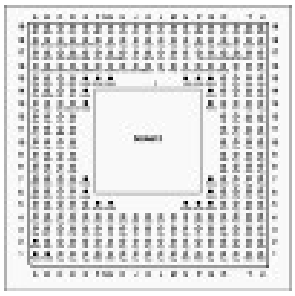


Figure: 238-pin Intel Socket 2 configuration.

Notice that although the chip for Socket 2 is called Pentium OverDrive, it is not a full-scale (64-bit) Pentium. Intel released the design of Socket 2 a little prematurely and found that the chip ran too hot for many systems. The company solved this problem by adding a special active heatsink to the Pentium OverDrive processor. This active heatsink is a combination of a standard heatsink and a built-in electric fan. Unlike the aftermarket glue-on or clip-on fans for processors that you might have seen, this one actually draws 5V power directly from the socket to drive the fan. No external connection to disk drive cables or the power supply is required. The fan/heatsink assembly clips and plugs directly into the processor and provides for easy replacement if the fan fails.

Another requirement of the active heatsink is additional clearance—no obstructions for an area about 1.4" off the base of the existing socket to allow for heatsink clearance. The Pentium OverDrive upgrade is difficult or impossible in systems that were not designed with this feature.

Another problem with this particular upgrade is power consumption. The 5V Pentium OverDrive processor draws up to 2.5 amps at 5V (including the fan) or 12.5 watts, which is more than double the 1.2 amps (6 watts) drawn by the DX2 66 processor.

Socket 3

Because of problems with the original Socket 2 specification and the enormous heat the 5V version of the Pentium OverDrive processor generates, Intel came up with an improved design. This processor is the same as the previous Pentium OverDrive processor, except that it runs on 3.3V and draws a maximum 3.0 amps of 3.3V (9.9 watts) and 0.2 amp of 5V (1 watt) to run the fan—a total of 10.9 watts. This configuration provides a slight margin over the 5V version of this processor. The fan is easy to remove from the OverDrive processor for replacement, should it ever fail.

Intel had to create a new socket to support both the DX4 processor, which runs on 3.3V, and the 3.3V Pentium OverDrive processor. In addition to the 3.3V chips, this new socket supports the older 5V SX, DX, DX2, and even the 5V Pentium OverDrive chip. The design, called Socket 3, is the most flexible upgradeable 486 design. Figure shows the pinout specification of Socket 3.

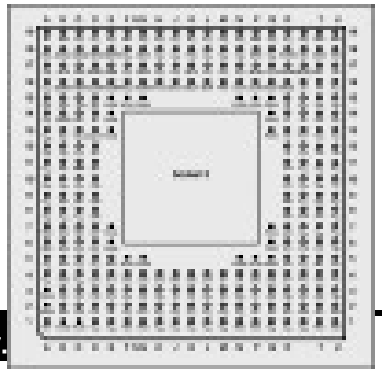


Figure. 237-pin Intel Socket 3 configuration.

Notice that Socket 3 has one additional pin and several others plugged in compared with Socket 2. Socket 3 provides for better keying, which prevents an end user from accidentally installing the processor in an improper orientation. However, one serious problem exists: This socket can't automatically determine the type of voltage that is provided to it. You will likely find a jumper on the motherboard near the socket to enable selecting 5V or 3.3V operation.

Caution

Because this jumper must be manually set, a user could install a 3.3V processor in this socket when it is configured for 5V operation. This installation instantly destroys the chip when the system is powered on. So, it is up to the end user to ensure that this socket is properly configured for voltage, depending on which type of processor is installed. If the jumper is set in 3.3V configuration and a 5V processor is installed, no harm will occur, but the system will not operate properly unless the jumper is reset for 5V.

Socket 4

Socket 4 is a 273-pin socket designed for the original Pentium processors. The original Pentium 60MHz and 66MHz version processors had 273 pins and plugged into Socket 4. It is a 5V-only socket because all the original Pentium processors run on 5V. This socket accepts the original Pentium 60MHz or 66MHz processor and the OverDrive processor. Figure . shows the pinout specification of Socket 4.

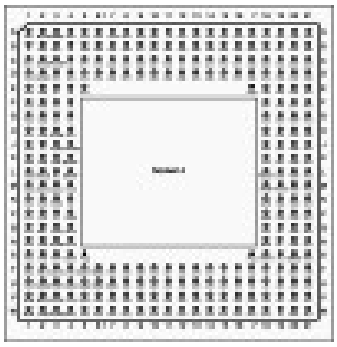


Figure . 273-pin Intel Socket 4 configuration.

Somewhat amazingly, the original Pentium 66MHz processor consumes up to 3.2 amps of 5V power (16 watts), not including power for a standard active heatsink (fan).

The 66MHz OverDrive processor that replaced it consumes a maximum 2.7 amps (13.5 watts), including about 1 watt to drive the fan. Even the original 60MHz Pentium processor consumes up to 2.91 amps at 5V (14.55 watts). It might seem strange that the replacement processor, which is twice as fast, consumes less power than the original, but this has to do with the manufacturing processes used for the original and OverDrive processors.

Although both processors run on 5V, the original Pentium processor was created with a circuit size of 0.8 micron, making that processor much more power-hungry than the 0.6-micron circuits used in the OverDrive and the other Pentium processors. Shrinking the circuit size is one of the best ways to decrease power consumption. Although the OverDrive processor for Pentium-based systems draws less power than the original processor, additional clearance might have to be allowed for the active heatsink assembly that is mounted on top. As in other OverDrive processors with built-in fans, the power to run the fan is drawn directly from the chip socket, so no separate power-supply connection is required. Also, the fan is easy to replace should it ever fail.

Socket 5

When Intel redesigned the Pentium processor to run at 75MHz, 90MHz, and 100MHz, the company went to a 0.6-micron manufacturing process and 3.3V operation. This change resulted in lower power consumption: only 3.25 amps at 3.3V (10.725 watts). Therefore, the 100MHz Pentium processor used far less power than even the original 60MHz version. This resulted in lower power consumption and enabled the extremely high clock rates without overheating.

The Pentium 75 and higher processors actually have 296 pins, although they plug into the official Intel Socket 5 design, which calls for a total of 320 pins. The additional pins are used by the Pentium OverDrive for Pentium processors. This socket has the 320 pins configured in a staggered PGA, in which the individual pins are staggered for tighter clearance.

Several OverDrive processors for existing Pentiums were available. These usually were later design chips with integral voltage regulators to enable operating on the higher voltages the older chips originally required. Intel no longer sells these; however, companies such as PowerLeap do still sell upgrade chips for older systems. Figure. shows the standard pinout for Socket 5.

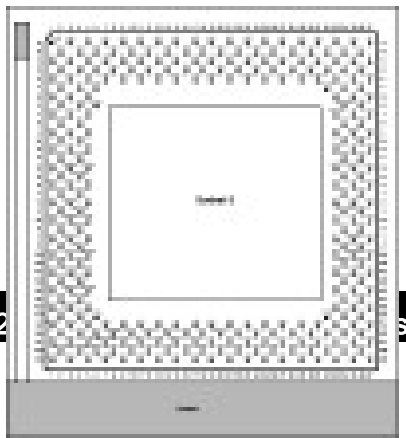


Figure. 320-pin Intel Socket 5 configuration.

The Pentium OverDrive for Pentium processors has an active heatsink (fan) assembly that draws power directly from the chip socket. The chip requires a maximum 4.33 amps of 3.3V to run the chip (14.289 watts) and 0.2 amp of 5V power to run the fan (one watt), which results in a total power consumption of 15.289 watts. This is less power than the original 66MHz Pentium processor requires, yet it runs a chip that is as much as four times faster!

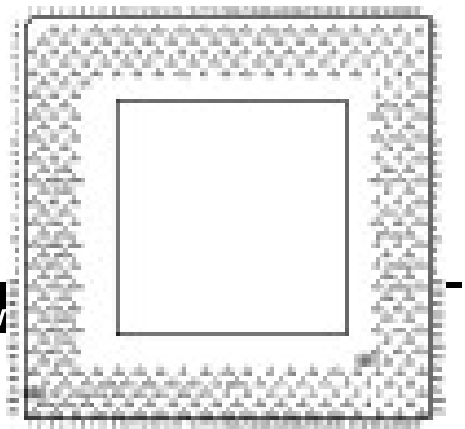
Socket 6

The last 486 socket was designed for the 486 DX4 and the 486 Pentium OverDrive processor. Socket 6 was intended as a slightly redesigned version of Socket 3 and had an additional 2 pins plugged for proper chip keying. Socket 6 has 235 pins and accepts only 3.3V 486 or OverDrive processors. Although Intel went to the trouble of designing this socket, it never was built or implemented in any systems. Motherboard manufacturers instead stuck with Socket 3.

Socket 7 (and Super7)

Socket 7 is essentially the same as Socket 5 with one additional key pin in the opposite inside corner of the existing key pin. Socket 7, therefore, has 321 pins total in a 37x37 SPGA arrangement. The real difference with Socket 7 is not with the socket itself, but with the companion voltage regulator module (VRM) circuitry on the motherboard that must accompany it.

The VRM is either a small circuit board or a group of circuitry embedded in the motherboard that supplies the proper voltage level and regulation of power to the processor.



The main reason for the VRM is that Intel and AMD wanted to drop the voltages the processors would use from the 3.3V or 5V supplied to the motherboard by the power supply. Rather than require custom power

supplies for different processors, the VRM converts the 3.3V or 5V to the proper voltage for the particular CPU you are using. Intel released different versions of the Pentium and Pentium-MMX processors that ran on 3.3V (called VR), 3.465V (called VRE), or 2.8V. Equivalent processors from AMD, Cyrix, and others used voltages from 3.3V to 1.8V. Because of the variety of voltages that might be required to support different processors, most motherboard manufacturers started including VRM sockets or building adaptable VRMs into their Pentium motherboards.

Figure. shows the Socket 7 pinout.

Figure. Socket 7 (Pentium) pinout (top view).

AMD, along with Cyrix and several chipset manufacturers, pioneered an improvement or extension to the Intel Socket 7 design called Super Socket 7 (or Super7), taking it from 66MHz to 95MHz and 100MHz. This enabled faster Socket 7–type systems to be made, supporting processors up to 500MHz, which are nearly as fast as some of the newer Slot 1– and Socket 370–type systems using Intel processors. Super7 systems also have support for the AGP video bus, as well as Ultra DMA hard disk controllers and advanced power management.

Major third-party chipset suppliers—including Acer Laboratories, Inc. (ALi); VIA Technologies; and Silicon Integrated Systems (SiS)—all released chipsets for Super7 boards. Most of the major motherboard manufacturers made Super7 boards in both Baby-AT and ATX form factors.

Socket 8

Socket 8 is a special SPGA socket featuring a whopping 387 pins! This was specifically designed for the Pentium Pro processor with the integrated L2 cache. The additional pins are required by the P6 processor bus. Figure. shows the Socket 8 pinout.

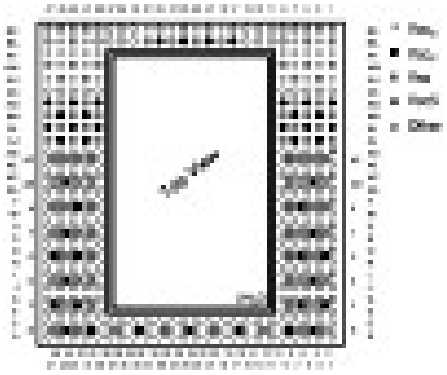


Figure. Socket 8 (Pentium Pro) pinout showing power pin locations.

Socket 370 (PGA-370)

In November 1998, Intel introduced a new socket for P6 class processors. The socket was called Socket 370 or PGA-370 because it has 370 pins and originally was designed for lower-cost PGA versions of the Celeron and Pentium III processors. Socket 370 was originally designed to directly compete in the lower-end system market along with the Super7 platform supported by AMD and Cyrix. However, Intel later used it for the Pentium III processor. Initially all the Celeron and Pentium III processors were made in SECC or SEPP format. These are essentially circuit boards containing the processor and separate L2 cache chips on a small board that plugs into the motherboard via Slot 1. This type of design was necessary when the L2 cache chips were made a part of the processor but were not directly integrated into the processor die. Intel did make a multiple-die chip package for the Pentium Pro, but this proved to be a very expensive way to package the chip, and a board with separate chips was cheaper, which is why the Pentium II looks different from the Pentium Pro.

Starting with the Celeron 300A processor introduced in August 1998, Intel began combining the L2 cache directly on the processor die; it was no longer in separate chips. With the cache fully integrated into the die, there was no longer a need for a board-mounted processor. Because it costs more to make a Slot 1 board or cartridge-type

processor instead of a socketed type, Intel moved back to the socket design to reduce the manufacturing cost—especially with the Celeron, which at that time was competing on the low end with Socket 7 chips from AMD and Cyrix.

The Socket 370 (PGA-370) pinout is shown in Figure ..

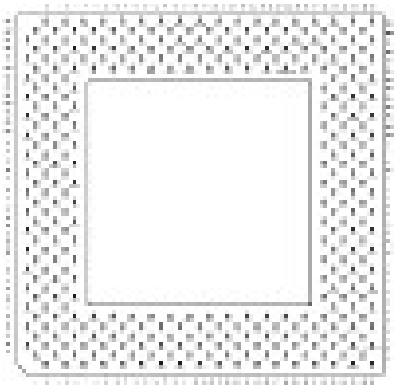


Figure. Socket 370 (PGA-370) Pentium III/Celeron pinout (top view).

The Celeron was gradually shifted over to PGA-370, although for a time both were available. All Celeron processors at 333MHz and lower were available only in the Slot 1 version. Celeron processors from 366MHz to 433MHz were available in both Slot 1 and Socket 370 versions; all Celeron processors from 466MHz and up through 1.4GHz are available only in the Socket 370 version.

Starting in October 1999, Intel also introduced Pentium III processors with integrated cache that plug into Socket 370. These use a packaging called flip chip pin grid array (FC-PGA), in which the raw die is mounted on the substrate upside down. The slot version of the Pentium III was more expensive and no longer necessary because of the on-die L2 cache.

Note that because of some voltage changes and one pin change, many original Socket 370 motherboards do not accept the later FC-PGA Socket 370 versions of the Pentium III and Celeron. Pentium III processors in the FC-PGA form have two RESET pins and require VRM 8.4 specifications. Prior motherboards designed only for the older versions of the Celeron are referred to as *legacy motherboards*, and the newer motherboards supporting the second RESET pin and VRM 8.4 specification are referred to as *flexible motherboards*. Contact your motherboard or system manufacturer for information to see whether your socket is the flexible version. Some motherboards, such as the Intel CA810, do support the VRM 8.4 specifications and supply proper voltage, but without Vtt support the Pentium III processor in the FC-PGA package will be held in RESET#. The last versions of the Pentium III and Celeron III use the Tualatin core design, which also requires a revised socket to operate. Motherboards that can handle Tualatin-core processors are known as *Tualatin-ready* and use different chipsets from those not designed to work with the Tualatin-core processor. Companies that sell upgrade processors offer products that enable you to install a Tualatin-core Pentium III or Celeron III processor into a motherboard that lacks built-in Tualatin support.

Installing a Pentium III processor in the FC-PGA package into an older motherboard is unlikely to damage the motherboard. However, the processor itself could be damaged. Pentium III processors in the 0.18-micron process operate at either 1.60V or 1.65V, whereas the Intel Celeron processors operate at 2.00V. The motherboard could be damaged if the motherboard BIOS fails to recognize the voltage identification of the processor. Contact your PC or motherboard manufacturer before installation to ensure compatibility.

A motherboard with a Slot 1 can be designed to accept almost any Celeron, Pentium II, or Pentium III processor. To use the socketed Celerons and Pentium III processors, several manufacturers have made available a low-cost slot-to-socket adapter sometimes called a *slot-ket*. This is essentially a Slot 1 board containing only a Socket 370, which enables you to use a PGA processor in any Slot 1 board. A typical slot-ket adapter is shown in the "Celeron" section later in this chapter.

Socket 423

Socket 423 is a ZIF-type socket introduced in November 2000 for the original Pentium 4. Figure . shows Socket 423.

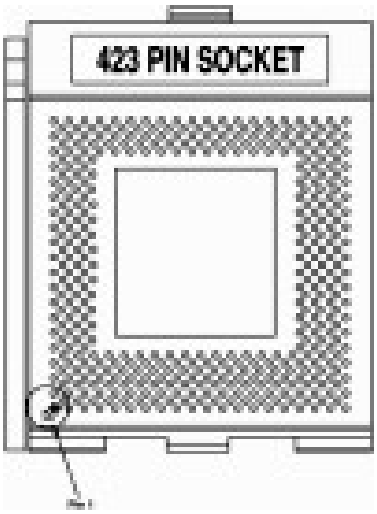


Figure. Socket 423 (Pentium 4) showing pin 1 location.

Socket 423 supports a 400MHz processor bus, which connects the processor to the Memory Controller Hub (MCH), which is the main part of the motherboard chipset and similar to the North Bridge in earlier chipsets. Pentium 4 processors up to 2GHz were available for Socket 423; all faster versions require Socket 478 instead.

Socket 423 uses a unique heatsink mounting method that requires standoffs attached either to the chassis or to a special plate that mounts underneath the motherboard. This was designed to support the weight of the larger heatsinks required for the Pentium 4. Because of this, many Socket 423 motherboards require a special chassis that has the necessary additional standoffs installed. Fortunately, the need for these standoffs was eliminated with the newer Socket 478 for Pentium 4 processors.

The processor uses five voltage ID (VID) pins to signal the VRM built into the motherboard to deliver the correct voltage for the particular CPU you install. This makes the voltage selection completely automatic and foolproof. Most Pentium 4 processors for Socket 423 require 1.7V. A small triangular mark indicates the pin-1 corner for proper orientation of the chip.

Socket 478

Socket 478 is a ZIF-type socket for the Pentium 4 and Celeron 4 (Celerons based on the Pentium 4 core) introduced in October 2001. It was specially designed to support additional pins for future Pentium 4 processors and speeds over 2GHz. The heatsink mounting is different from the previous Socket 423, allowing larger heatsinks to be attached to the CPU. Figure. shows Socket 478.

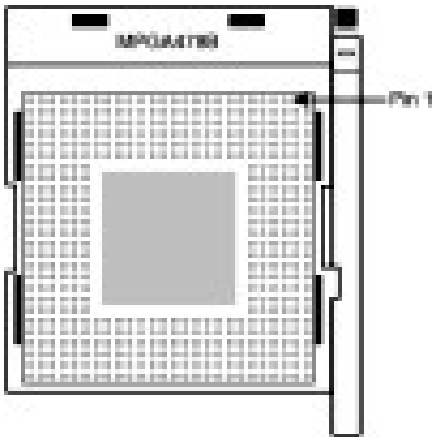
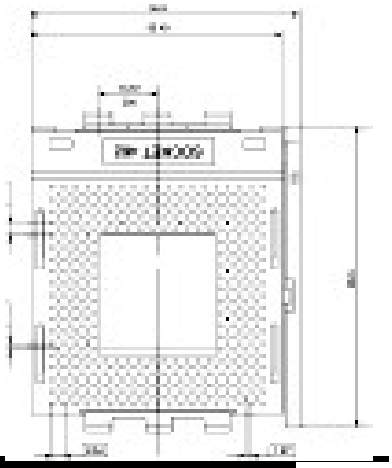


Figure. Socket 478 (Pentium 4) showing pin 1 location.

Socket 478 supports a 400MHz, 533MHz, or 800MHz processor bus that connects the processor to the memory controller hub (MCH), which is the main part of the motherboard chipset.

Socket 478 uses a heatsink attachment method that clips the heatsink directly to the motherboard, and not the CPU socket or chassis (as with Socket 423). Therefore, any standard chassis can be used, and the special standoffs used by Socket 423 boards are not required. This heatsink attachment allows for a much greater clamping load between the heatsink and processor, which aids cooling.

Socket 478 processors use five VID pins to signal the VRM built into the motherboard to deliver the correct voltage for the particular CPU you install. This makes the voltage selection completely automatic and foolproof. A small triangular mark indicates the pin-1 corner for proper orientation of the chip.



Socket A (Socket 462)

AMD introduced Socket A, also called Socket 462, in June 2000 to support the PGA versions of the Athlon and Duron processors. It is designed as a replacement for Slot A used by the original Athlon processor. Because the Athlon has now

moved to incorporate L2 cache on-die, and the low-cost Duron was manufactured only in an on-die cache version, there was no longer a need for the expensive cartridge packaging the original Athlon processors used.

Socket A has 462 pins and 11 plugs oriented in an SPGA form (see Figure). Socket A has the same physical dimensions and layout as Socket 370; however, the location and placement of the plugs prevent Socket 370 processors from being inserted. Socket A supports 31 voltage levels from 1.100V to 1.850V in 0.025V increments, controlled by the VID0-VID4 pins on the processor. The automatic voltage regulator module circuitry typically is embedded on the motherboard.

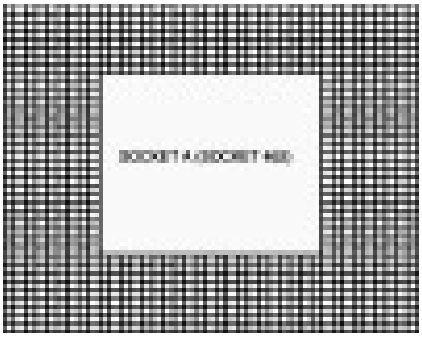


Figure. Socket A (Socket 462) Athlon/Duron layout.

There are 11 total plugged holes, including 2 of the outside pin holes at A1 and AN1. These are used to allow for keying to force the proper orientation of the processor in the socket. The pinout of Socket A is shown in Figure.

Figure. Socket A (Socket 462) Athlon/Duron pinout (top view).

After the introduction of Socket A, AMD moved all Athlon (including all Athlon XP) processors to this form factor, phasing out Slot A. In addition, for a time AMD also sold a reduced L2 cache version of the Athlon called the Duron in this form factor. In 2005, AMD discontinued the Athlon XP and introduced the AMD Sempron in both Socket A and Socket 754 form factors. The first Athlon 64 processors also used Socket 754, but most current Athlon 64 processors now use Socket 939.

Caution

Just because a chip can plug into a socket doesn't mean it will work. The newer Athlon XP and Socket A Sempron processors require different voltages, BIOS, and chipset support than earlier Socket A Athlon and Duron processors. As always, make sure your motherboard supports the processor you intend to install.

Socket 603

Socket 603 is used with the Intel Xeon processor in DP (dual processor) and MP (multiple processor) configurations. These are typically used in motherboards designed for use in network file servers. Figure shows Socket 603.

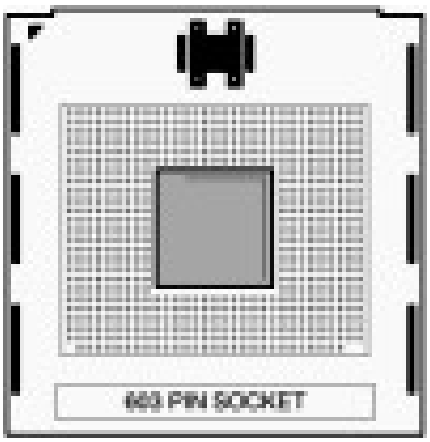


Figure 3.25 Socket 603 is used by the Intel Xeon processor.

Socket 754

Socket 754 is used with the initial releases of the AMD Athlon 64 processors. Socket 754 is also used by some versions of the AMD Sempron, AMD's economy processor line. This socket supports single-channel unbuffered DDR SDRAM. Figure 3.26 shows an overhead view of this socket.

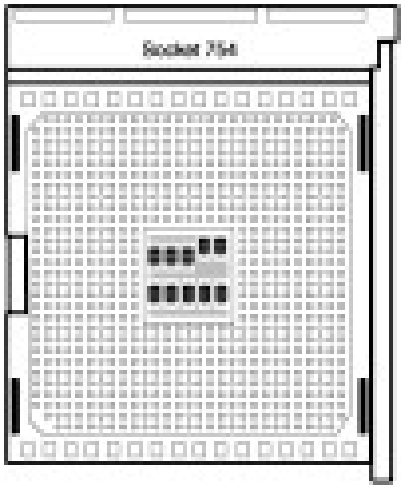


Figure 3.26 Socket 754. The large cutout corner at the lower left indicates pin 1.

Socket 939 and 940

Socket 939 is used with the Socket 939 versions of the AMD Athlon 64, 64 FX, and 64 X2 (see Figure). It's also used by some recent versions of the AMD Opteron processor for workstations and servers. Motherboards using this socket support conventional unbuffered DDR SDRAM modules in either single- or dual-channel mode, rather than the server-oriented (more expensive) registered modules required by Socket 940 motherboards. Sockets 939 and 940 have different pin arrangements and processors for each and are not interchangeable.

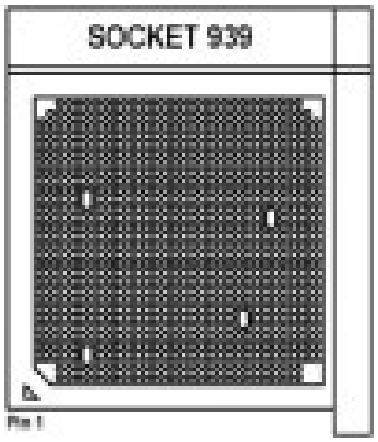


Figure .Socket 939. The cutout corner and triangle at the lower left indicate pin 1.

Socket 940 is used with the Socket 940 version of the AMD Athlon 64 FX, as well as most AMD Opteron processors (see Figure). Motherboards using this socket support only registered DDR SDRAM modules in dual-channel mode. Because the pin arrangement is different, Socket 939 processors do not work in Socket 940, and vice versa.

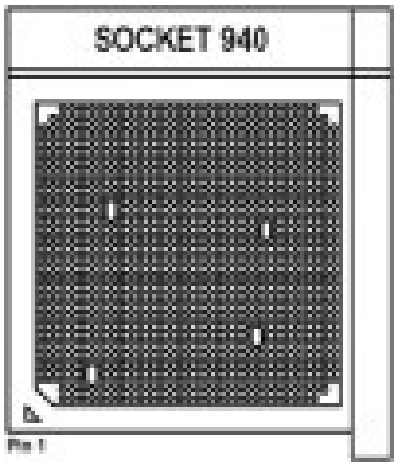


Figure. Socket 940. The cutout corner and triangle at the lower left indicate pin 1.

Socket T

Socket T (LGA775) is used by the latest versions of the Intel Pentium 4 Prescott processor and the Pentium D and Pentium Extreme Edition processors, as well as some versions of the Celeron D. The first-generation Prescott processors used Socket 478. Socket T is unique in that it uses a land grid array format, so the pins are on the socket, rather than the processor. The first LGA processors were the Pentium II and Celeron processors in 1997; in those processors LGA packaging was used for the chip mounted on the Slot-1 cartridge.

LGA uses gold pads (called *lands*) on the bottom of the substrate to replace the pins used in PGA packages. In socketed form, it allows for much greater clamping forces and therefore greater stability and improved thermal transfer (better cooling). LGA is really just a recycled version of what was previously called LCC (leadless chip carrier) packaging. This was used way back on the 286 processor in '84, which had gold lands around the edge only (there were far fewer pins back then). In other ways LGA is simply a modified version of ball grid array (BGA), with gold lands replacing the solder balls, making it more suitable for socketed (rather than soldered) applications. The early LCC packages were ceramic, whereas the first Pentium II LGA packages were plastic, with the package soldered to a cartridge substrate. These days (and for the future) the LGA package is organic and directly socketed instead. On a technical level, the Pentium 4 LGA chips combine several packaging technologies that have all been used in the past, including organic land grid array (OLGA) for the substrate and controlled collapse chip connection (C4) flip-chip for the actual processor die (see Figure).

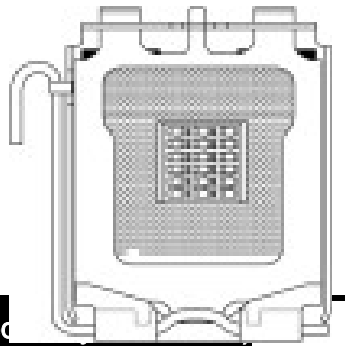


Figure. Socket T. The release lever on the left is used to raise the clamp out of the way to permit the processor to be placed over the contacts.

Socket M2

In the second quarter of 2006, AMD introduced processors that use a new socket, called Socket M2 (see Figure). AMD intends for M2 to be the eventual replacement for the confusing array of Socket 754, Socket 939, and Socket 940 form factors it uses for the Athlon 64, Athlon 64 FX, Athlon 64 X2, Opteron, and Socket 754 AMD Sempron processors.

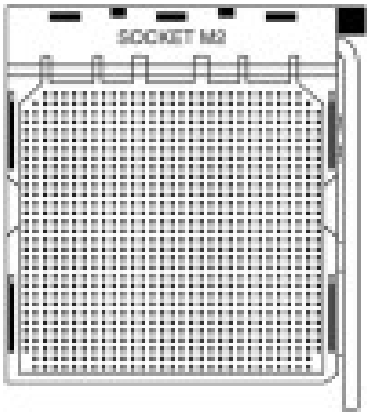


Figure. Socket M2. The cutout corner at the lower left indicates pin 1.

Although Socket M2 contains 940 pins—the same number as used by Socket 940—Socket M2 is designed to support the integrated dual-channel DDR2 memory controllers that were added to the Athlon 64 and Opteron processor families in 2006. Processors designed for Sockets 754, 939, and 940 include DDR memory controllers and are not pin compatible with Socket M2.

Processor Slots

After introducing the Pentium Pro with its integrated L2 cache, Intel discovered that the physical package it chose was very costly to produce. Intel was looking for a way to easily integrate cache and possibly other components into a processor package, and it came up with a cartridge or board design as the best way to do this. To accept its new cartridges, Intel designed two types of slots that could be used on motherboards.

Slot 1 is a 242-pin slot designed to accept Pentium II, Pentium III, and most Celeron processors. Slot 2, on the other hand, is a more sophisticated 330-pin slot designed for the Pentium II Xeon and Pentium III Xeon processors, which are primarily for workstations and servers. Besides the extra pins, the biggest difference between Slot 1 and Slot 2 is the fact that Slot 2 was designed to host up to four-way or more processing in a single board. Slot 1 allows only single or dual processing functionality.

Note that Slot 2 is also called SC330, which stands for slot connector with 330 pins. Intel later discovered less-expensive ways to integrate L2 cache into the processor core and no longer produces Slot 1 or Slot 2 processors. Both Slot 1 and Slot 2 processors are now obsolete, and many systems using these processors have been retired or upgraded with socket-based motherboards.

Slot 1 (SC242)

Slot 1, also called SC242 (slot connector 242 pins), is used by the SEC design that is used with the cartridge-type Pentium II/III and Celeron processors (see Figure).

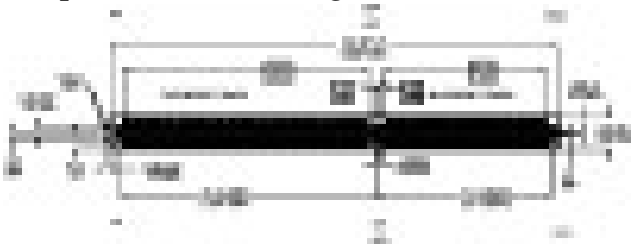


Figure. Slot 1 connector dimensions and pin layout.

Slot 2 (SC330)

Slot 2, otherwise called SC330 (slot connector 330 pins), is used on high-end motherboards that support the Pentium II Xeon and Pentium III Xeon processors. Figure. shows the Slot 2 connector.

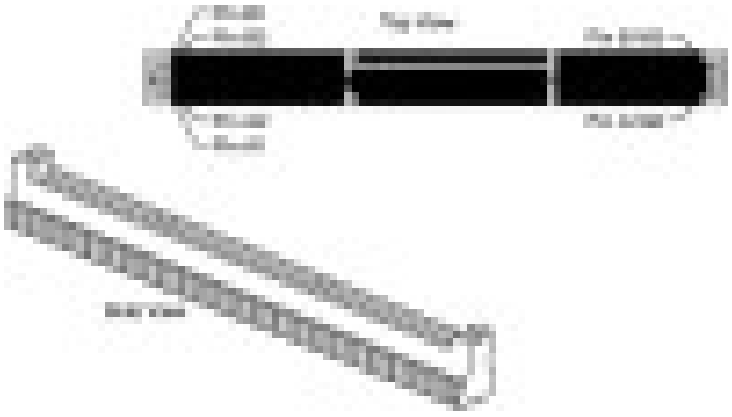


Figure. Slot 2 (SC330) connector dimensions and pin layout.

The Pentium II Xeon and Pentium III Xeon processors are designed in a cartridge similar to, but larger than, that used for the standard Pentium II/III. Figure. shows the Xeon cartridge.

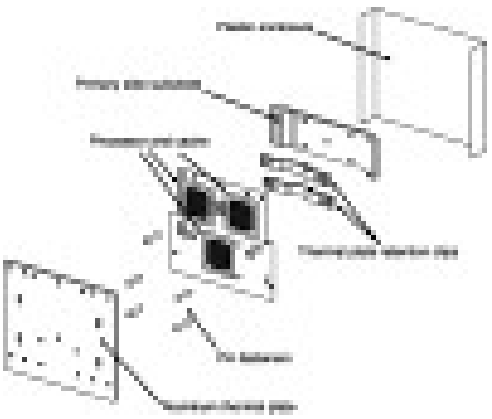


Figure 3.33 Pentium II/III Xeon cartridge.

Slot 2 motherboards were used in higher-end systems such as workstations or servers based on the Pentium II Xeon or Pentium III Xeon. These versions of the Xeon differ from the standard Pentium II and slot-based Pentium III mainly by virtue of having full-core speed L2 cache, and in some versions more of it. The additional pins allow for additional signals needed by multiple processors.

Table 3.18. CPU Socket and Slot Types and Specifications

Chip Class	Socket	Pins	Layout	Voltage	Supported Processors	Introduced
Intel/AMD 486 class	Socket 1	169	17x17 PGA	5V	486 SX/SX2, DX/DX2, DX4 OD	Apr. '89
	Socket 2	238	19x19 PGA	5V	486 SX/SX2, DX/DX2, DX4 OD, 486 Pentium OD	Mar. '92
	Socket 3	237	19x19 PGA	5V/3.3V	486 SX/SX2, DX/DX2, DX4, 486 Pentium OD, AMD 5x86	Feb. '94
	Socket 6 [1]	235	19x19 PGA	3.3V	486 DX4, 486 Pentium OD	Feb. '94
Intel/AMD 586 (Pentium) class	Socket 4	273	21x21 PGA	5V	Pentium 60/66, OD	Mar. '93
	Socket 5	320	37x37 SPGA	3.3V/3.5V	Pentium 75-133, OD	Mar. '94
	Socket 7	321	37x37 SPGA	VRM	Pentium 75-233+, MMX, OD, AMD K5/K6, Cyrix M1/II	June '95
Intel 686 (Pentium II/III) SPGA	Socket 8	387	Dual-pattern SPGA	Auto VRM	Pentium Pro, OD	Nov. '95
class	Slot 1 (SC242)	242	Slot	Auto VRM	Pentium II/III, Celeron SECC	May '97
	Socket 370	370	37x37 SPGA	Auto VRM	Celeron/Pentium III PPGA/FC-PGA	Nov. '98
Intel Pentium 4 class	Socket 423	423	39x39 SPGA	Auto VRM	Pentium 4 FC-PGA	Nov. '00
	Socket 478	478	26x26 mPGA	Auto VRM	Pentium 4/Celeron FC-PGA2	Oct. '01
	Socket T (LGA775)	775	30x33 LGA	Auto VRM	Pentium 4/Celeron/Pentium	June '04

Chip Class	Socket	Pins	Layout	Voltage	Supported Processors	Introduced
					D/Pentium Extreme Edition/LGA775	
AMD K7 class	Slot A	242	Slot	Auto VRM	AMD Athlon SECC	June '99
	Socket A (462)	462	37x37 SPGA	Auto VRM	AMD Athlon/Athlon XP/Duron PGA/FC-PGA	June '00
AMD K8 class	Socket 754	754	29x29 mPGA	Auto VRM	AMD Athlon 64	Sep. '03
	Socket 939	939	31x31 mPGA	Auto VRM	AMD Athlon 64 v.2	June '04
	Socket 940	940	31x31 mPGA	Auto VRM	AMD Athlon 64FX, Opteron	Apr. '03
Intel/AMD server and workstation class	Slot 2(SC330)	330	Slot	Auto VRM	Pentium II/III Xeon	Apr. '98
	Socket 603	603	31x25 mPGA	Auto VRM	Xeon (P4)	May '01
	Socket 604	604	31x25 mPGA	Auto VRM	Xeon (P4)	Oct. '03
	Socket PAC418	418	38x22	Auto VRM split SPGA	Itanium	May '01
	Socket PAC611	611	25x28	Auto VRM mPGA	Itanium 2	July '02
	Socket 940	940	31x31 mPGA	Auto VRM	AMD Athlon 64FX, Opteron	Apr. '03
FC-PGA = Flip-chip pin grid array						
FC-PGA2 = FC-PGA with an Integrated Heat Spreader (IHS)						
OD = OverDrive (retail upgrade processors)						
PAC = Pin array cartridge						

Chip Class	Socket	Pins	Layout	Voltage	Supported Processors	Introduced
PGA = Pin grid array						
PPGA = Plastic pin grid array						
SC242 = Slot connector, 242 pins						
SC330 = Slot connector, 330 pins						
SECC = Single edge contact cartridge						
SPGA = Staggered pin grid array						
mPGA = Micro pin grid array						
VRM = Voltage regulator module with variable voltage output determined by module type or manual jumpers						
Auto VRM = Voltage regulator module with automatic voltage selection determined by processor Voltage ID (VID) pins						

Reference : http://www.quepublishing.com/articles/article.aspx?p=482324&seqNum=6

ChipSelect Basic

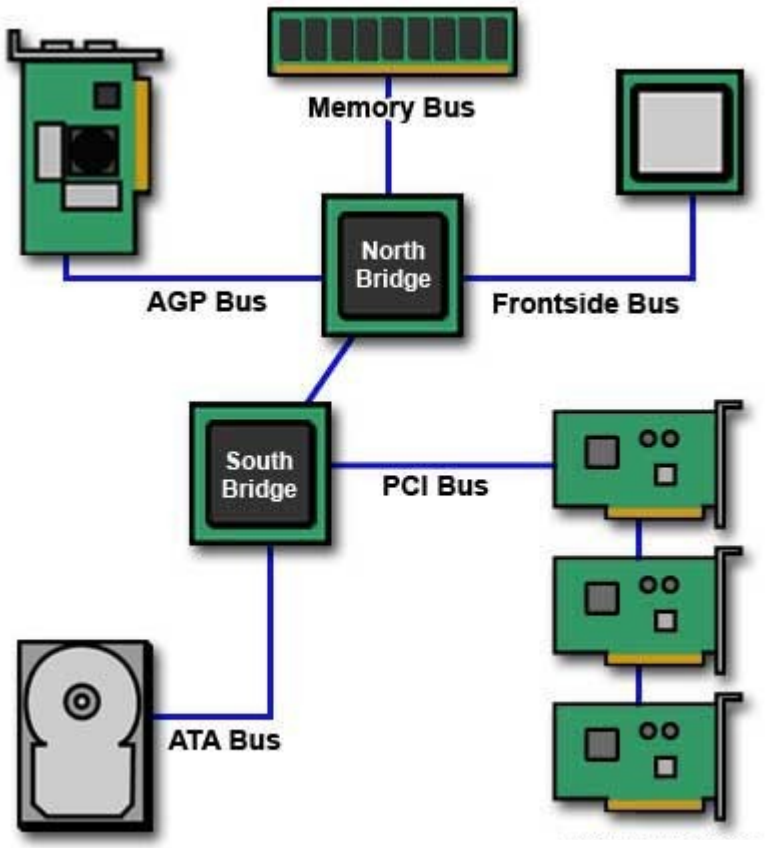
Q.What is Chipset?

Ans. A number of integrated circuits designed to perform one or more related functions. For example, one chipset may provide the basic functions of a modem while another provides the CPU functions for a computer. Newer chipsets generally include functions provided by two or more older chipsets. In some cases, older chipsets that required two or more physical chips can be replaced with a chipset on one chip. The term is often used to refer to the core functionality of a motherboard.

On a PC, it consists of two basic parts -- the **northbridge** and the **southbridge**. All of the various components of the computer communicate with the CPU through the chipset.

The northbridge connects directly to the processor via the front side bus (FSB). A memory controller is located on the northbridge, which gives the CPU fast access to the memory. The northbridge also connects to the AGP or PCI Express bus and to the memory itself.

The southbridge is slower than the northbridge, and information from the CPU has to go through the northbridge before reaching the southbridge. Other busses connect the southbridge to the PCI bus, the USB ports and the IDE or SATA hard disk connections.



Chipset selection and CPU selection go hand in hand, because manufacturers optimize chipsets to work with specific CPUs. The chipset is an integrated part of the motherboard, so it cannot be removed or upgraded. This means that not only must the motherboard's socket fit the CPU, the motherboard's chipset must work optimally with the CPU.

Purpose of Chipset

A bus is simply a circuit that connects one part of the motherboard to another. The more data a bus can handle at one time, the faster it allows information to travel. The **speed** of the bus, measured in megahertz (MHz), refers to how much data can move across the bus simultaneously.

Bus speed usually refers to the speed of the **front side bus** (FSB), which connects the CPU to the northbridge. FSB speeds can range from 66 MHz to over 800 MHz. Since the CPU reaches the memory controller though the northbridge, FSB speed can dramatically affect a computer's performance.

Here are some of the other busses found on a motherboard:

- The **back side bus** connects the CPU with the level 2 (L2) cache, also known as secondary or external cache. The processor determines the speed of the back side bus.
- The **memory bus** connects the northbridge to the memory.

- The **IDE** or **ATA** bus connects the southbridge to the disk drives.
- The **AGP** bus connects the video card to the memory and the CPU. The speed of the AGP bus is usually 66 MHz.
- The **PCI** bus connects PCI slots to the southbridge. On most systems, the speed of the PCI bus is 33 MHz. Also compatible with PCI is **PCI Express**, which is much faster than PCI but is still compatible with current software and operating systems. PCI Express is likely to replace both PCI and AGP busses.

The faster a computer's bus speed, the faster it will operate -- to a point. A fast bus speed cannot make up for a slow processor or chipset.

There are two type of chipset architecture

1.Hub Architecture

2.Bridge architecture

http://computer.howstuffworks.com/motherboard4.htm

Q.Describe HUB architecture

Ans.

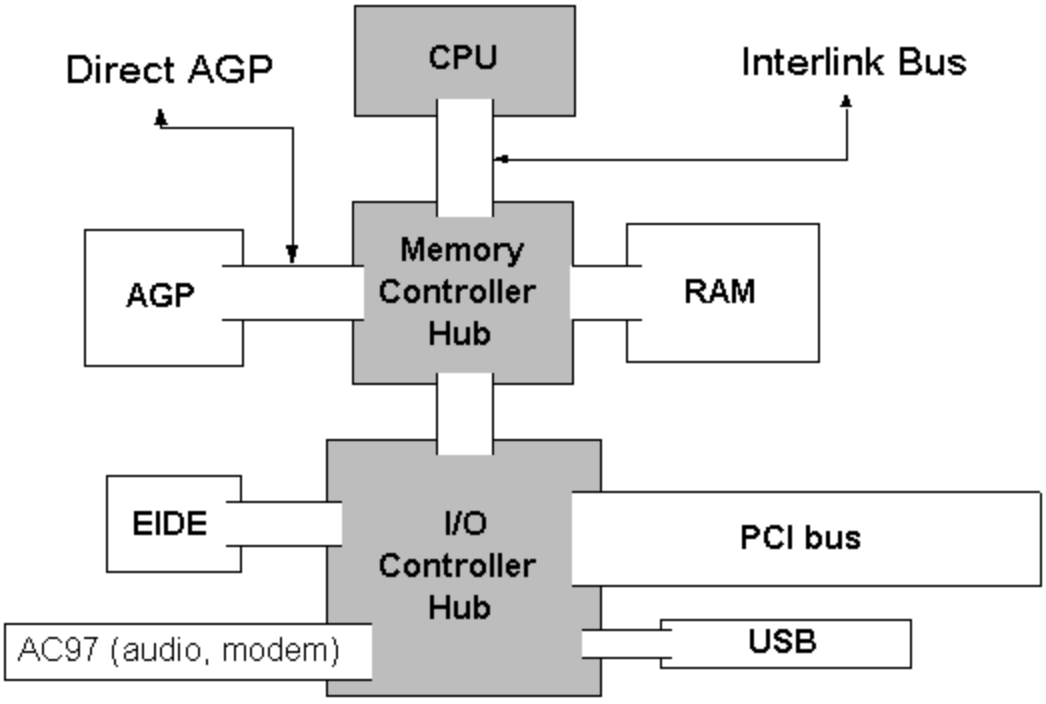


Fig.

Intel Hub Architecture (also called as AHA - *Accelerated Hub Architecture)*

Intel introduced this hub architecture starting with the 820 chipset. The hub architecture divides control between a memory controller hub (MCH) that supports memory and AGP and an I/O controller hub (ICH) that supports PCI, USB, sound, IDE and LAN. The word hub in Intel Hub Architecture refers to the north and south bridges in a chipset. **Intel has replaced those two terms with the word hub**.

Intel's architecture for the 8xx family of chipsets, starting with the 820. It uses a memory controller hub (MCH) that is connected to an I/O controller hub (ICH) via a 266 MB/sec bus. The MCH chip supports memory and AGP, while the ICH chip provides connectivity for PCI, USB, sound, IDE hard disks and LAN.

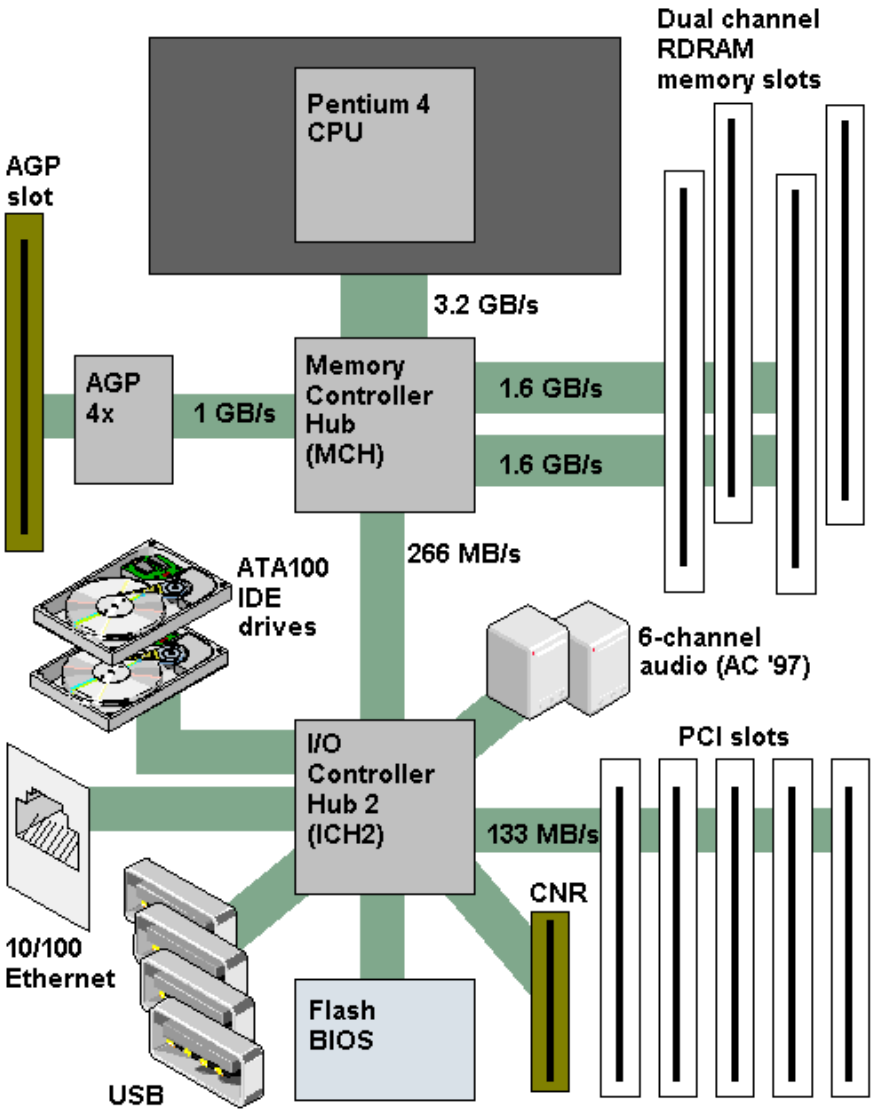


Fig.

Because of the high-speed channel between the sections, the Intel Hub Architecture (IHA) is much faster than the earlier Northbridge/Southbridge design, which hooked all low-speed ports to the PCI bus. The IHA also optimizes data transfer based on data type.

Accelerated Hub Architecture (AHA) (also called Intel Hub Architecture) is an Intel 800-series chipset design that uses a dedicated bus to transfer data between the two mainprocessor chips instead of using the Peripheral Component Interconnect (PCI) bus, which was used in previous chipset architectures. The Accelerated Hub Architecture provides twice the bandwidth of the traditional PCI bus architecture at 266 MB per second. The Accelerated Hub Architecture consists of a memory controller hub and an input/output (I/O) controller hub (a controller directs or manages access to devices).

The memory controller hub provides the central processing unit (CPU) interface, the memory interface, and the accelerated graphics port (AGP) interface. The memory controller hub supports single or dual processors with up to 1 GB of memory. The memory controller hub also allows for simultaneous processing, which enables more life-like audio and video capabilities.

The I/O controller hub provides a direct connection from the memory to the I/O devices, which includes any built-in modem and audio controllers, hard drives, Universal Serial Bus (USB) ports, and PCI add-in cards. The I/O controller hub also includes the Alert on LAN (local area network) feature that sounds an alert when software failures or system intrusion occurs.

http://www.techwarelabs.com/reviews/motherboard/albatron_px845pev/

Example

82810	Graphics Memory Controller Hub	421 Ball Grid Array (BGA)
82801	Integrated Controller Hub	241 Ball Grid Array (BGA)
82802	Firmware Hub	32-pin PLCC or 40-pin TSOP

82810 Graphics Memory Controller Hub

The 82810 Graphics Memory Controller Hub (GMCH) is a MCH "north bridge" including a graphics controller and using Direct AGP (integrated AGP, where the graphics controller is directly connected to the system RAM) operating at 100 MHz.

The 82810 chip features a "Hardware Motion Compensation" to improve soft DVD video and digital video out port for digital flat panel monitors. The graphics controller is a version of Intel's new model 752. Optional, the chip set can be equipped with a display cache of 4MB RAM to be used for "Z-buffering".

Dynamic Video Memory Technology (D.V.M.T.) is an architecture that offers good performance for the Value PC segment through efficient memory utilization and "Direct AGP". A new improved version of the SMBA (*Shared Memory Buffer Architecture*)used in earlier chip sets as VX. In the 810 chip set 11 MB system RAM is allocated to be used by the 3D-graphics controller as frame buffer, command buffer and Z-buffer.

82801 I/O Controller Hub

This "south bridge", the 82801 (ICH), employs an accelerated hub to give a direct connection from the graphics and memory to the integrated AC97 (Audio-Codec) controller, the IDE controllers, the dual USB ports, and the PCI bus. This promises increased I/O performance.

82802 Firmware Hub (FWH)

The 82802 Firmware Hub (FWH) stores system BIOS and video BIOS in a 4 Mbit EEPROM. In addition, the 82802 contains a hardware Random Number Generator (RNG), which (perhaps and in time) will enable better security, stronger encryption, and digital signing in the Internet.

AC97

The Integrated Audio-Codec 97 controller enables software audio and modem by using the processor to run sound and modem software. It will require software, but using this you need no modem or soundcard.

This feature is smart if you do not use audio or modem on a regular basis. It adds a heavy work to the CPU, which has to act as a modem and as a sound card beside its regular tasks.

Q.State Function of North and South Bridge

Ans. South Bridge

The **southbridge** is one of the two chips in the core logic chipset on a personal computer (PC) motherboard, the other being the northbridge. The southbridge typically implements the slower capabilities of the motherboard in a northbridge/southbridge chipset computer architecture. In Intel chipset systems, the southbridge is named *Input/Output Controller Hub (ICH)*. AMD, beginning with its Fusion APUs, has given the label *FCH*, or *Fusion Controller Hub*, to its southbridge.

The southbridge can usually be distinguished from the northbridge by not being directly connected to the CPU. Rather, the northbridge ties the southbridge to the CPU. Through the use of controller integrated channel circuitry, the northbridge can directly link signals from the I/O units to the CPU for data control and access.

Function of South Bridge

The south bridge is a chip on the motherboard. If we want to look at one of the latest models, we could take the south bridge, which is designed for motherboards with Pentium 4 processors. The south bridge incorporates a number of different controller functions, it looks after the transfer of data to and from the hard disk and all the other I/0 devices, and passes this data into the link channel which connects to the north bridge. It contains the following components and functions as shown in Table.

Component	Description
DMI	The south bridge is a chip on the motherboard. If we want to look at one of the latest models, we could take the south bridge, which is designed for motherboards with Pentium 4 processors. The south bridge incorporates a number of different controller functions, it looks after the transfer of data to and from the hard disk and all the other 1/0 devices, and passes this data into the link channel which connects to the north bridge. It contains the following components and functions as shown in Table.
PCi-Express	Hi-speed bus for I/O adapters.
PCi port	Standard I/O bus.
Serial Sata	Controller for up to four SATA hard disks.
Matrix Storage	Advanced Host Controller Interface for RAID0 and 1 on the same drives. Including support for Native Command Queuing and hot plug drive swaps.
Ultra Ata/100	Controller for PATA devices like hard disks, DVI and CD-

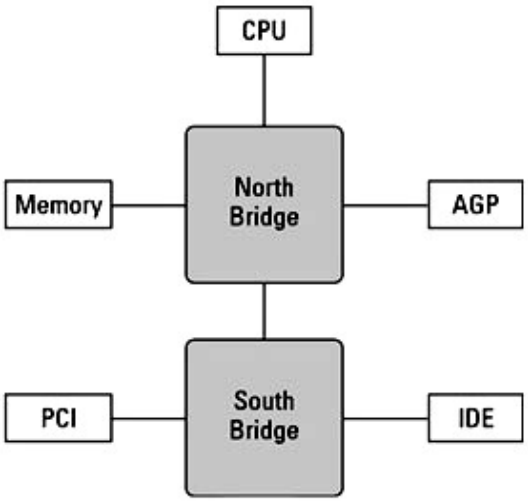
	drives.
USB Port	Hi-speed (JSB 2.0 ports.
7.1 Channel audio	Option for integrated sound device with Dolby Digital and UTS.
AC97 Modem	Integrated modem.
Ehternet	Integrated 10/100 Mbs network controller.

North Bridge

The **northbridge** or **host bridge** was one of the two chips in the core logic chipset on a PC motherboard, used to managedata communications between a CPU and a motherboard. It is supposed to be paired with a second support chip known as a southbridge.

The northbridge was historically one of the two chips in the core logic chipset on a PC motherboard, the other being thesouthbridge. Increasingly these functions became integrated into the CPU chip itself, beginning with memory and graphics controllers. For Intel Sandy Bridge and AMD Accelerated Processing Unit processors introduced in 2011, all of the functions of the northbridge reside on the CPU. When a separate northbridge is employed in older Intel systems, it is named *memory controller hub* (*MCH*) or *integrated memory controller hub* (*IMCH*) if equipped with an integrated VGA.

Function of North bridge



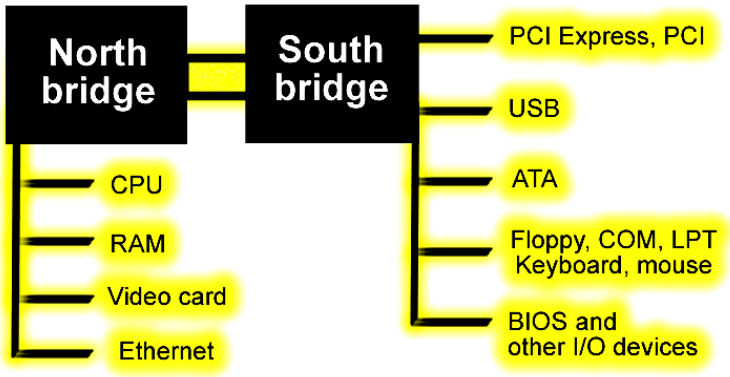


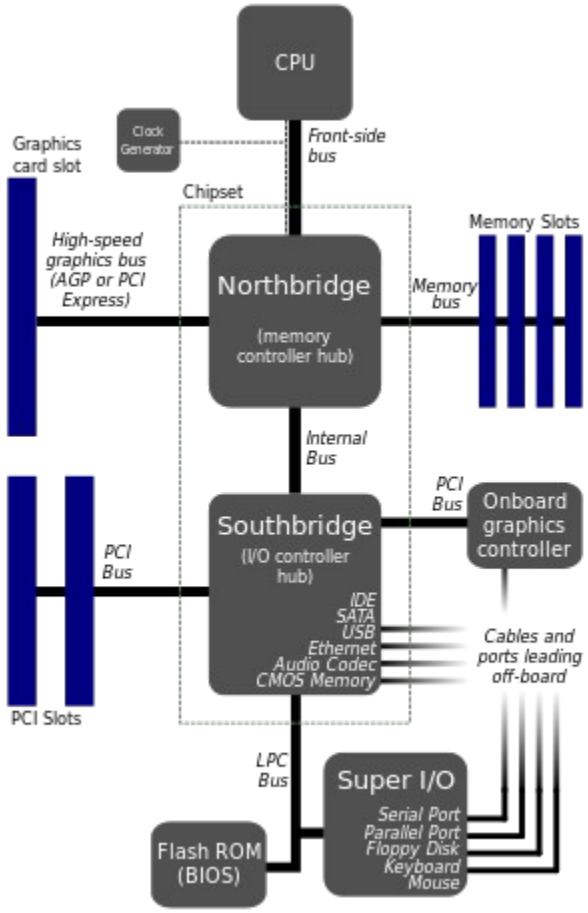
Fig. North and South Bridge – Bridge Architecture

The north bridge is a controller which controls the flow of data between the CPU and RAM, and to the AGP port. The north bridge has a large heat sink attached to it. It gets hot because of the often very large amounts of data traffic which pass it. The AGP is actually an I/0 port. It is used for the video card. In

contrast to the other I/O devices, the AGP port is connected directly to the north bridge, because it has to be as close to the RAM as possible. The same goes for the PC Express x16 port, which is the replacement of AGP in new motherboards.

Q.Why the name given North and South Bridge

Ans. The name is derived from drawing the architecture in the fashion of a map. The CPU would be at the top of the map comparable to due north on most general purpose geographical maps. The CPU would be connected to the chipset via a fast bridge (the northbridge) located north of other system devices as drawn. The northbridge would then be connected to the rest of the chipset via a slow bridge (the southbridge) located south of other system devices as drawn.



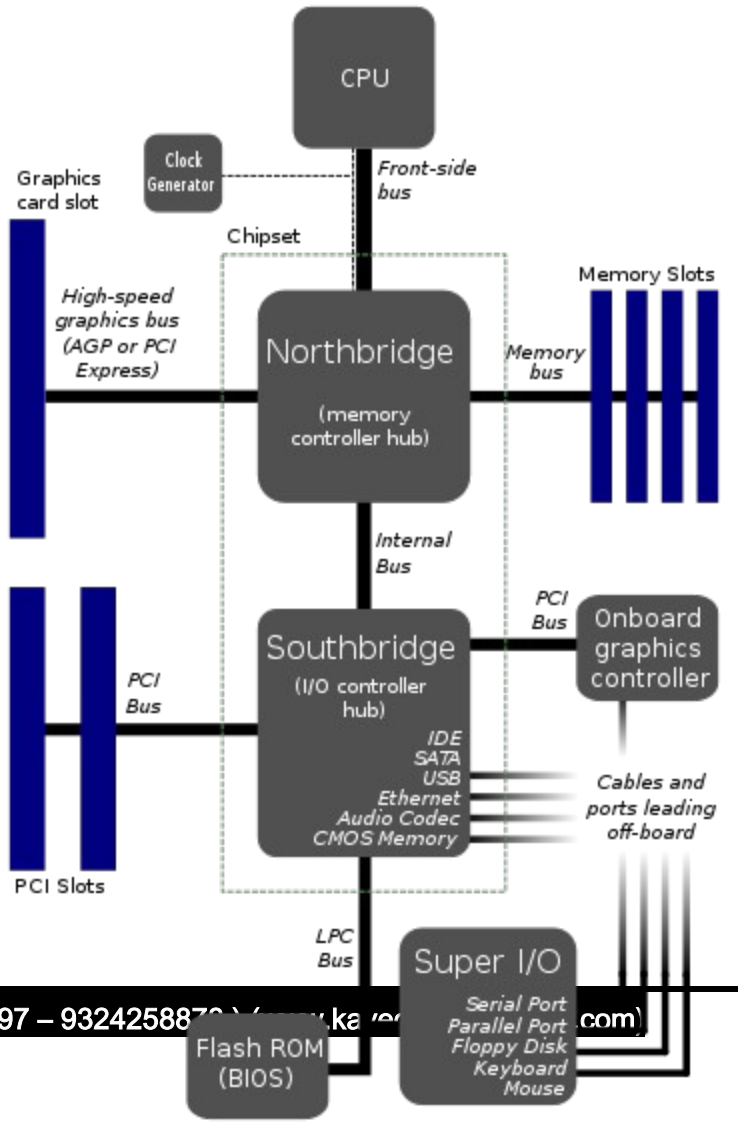
Q.How North bridge plays important role in over clocking

Ans. The northbridge plays an important part in how far a computer can be overclocked, as its frequency is commonly used as a baseline for the CPU to establish its own operating frequency. This chip typically gets hotter as processor speed becomes faster, requiring more cooling. There is a limit to CPU overclocking, as digital circuits are limited by physical factors such as propagation delay which increases with

(among other factors) operating temperature; consequently most overclocking applications have software-imposed limits on the multiplier and external clock setting.

Q.Describe the Chipset Architecture and State its Function

- ✓ **Ans.** The goals and needs of today's computer hardware customer are more diverse than ever before. Some people have a need for speed. Some will not buy a motherboard unless it has a list of specific features that he or she believes will be required for future upgrades. Others shy away from the cutting edge, instead requiring time-tested stability in a motherboard. Whether you are an over-clocker trying to squeeze the last M out of a CPU, or an IS manager who is looking for a corporate motherboard, you have to understand a motherboard's chipset
- ✓ A motherboard chipset has both a general definition and a specific definition that varies by chipset manufacturer. Generally speaking, a motherboard chipset controls the features and abilities of the motherboard. if you understand which chipset a motherboard uses, you know a good deal about its potential features and abilities before ever reading the motherboard's specifications.
- ✓ Modern motherboard chipsets nearly always consist of two separate chips. These two chips on the motherboard are called the north bridge and the south bridge. Together, the north bridge and the south bridge handle all of the communication between the processor, RAM, video options, PCI slots, BIOS, ATA controller, USB ports, integrated modem, integrated LAN port and integrated sound. The chipset also determines the type of RAM that can be used.
- ✓ There are a dozen or so reputable motherboard manufacturers and about a half dozen popular chipset



manufacturers. Intel and AMD provide specifications to the chipset manufacturers, who, in turn, develop and sell chipsets with various features and abilities to motherboard manufacturers. Of course, the exceptions to this are Intel and AMD, who also create their own chipsets.

Chipset architecture are of Two Types

- Hub Architecture
- North/South Architecture

North Bridge Architecture

Northbridge is an Intel chipset that communicates with the computer processor and controls interaction with memory, the Peripheral Component Interconnect (PCI) bus, Level 2 cache, and all Accelerated Graphics Port (AGP) activities. Northbridge communicates with the processor using the frontside bus (FSB). Northbridge is one part of a two-part chipset called Northbridge/Southbridge. Southbridge handles the input/output (I/O) functions of the chipset.

South Bridge Architecture

Southbridge is an Intel chipset that manages the basic forms of input/output (I/O) such as Universal Serial Bus (USB), serial , audio, Integrated Drive Electronics (IDE), and Industry Standard Architecture (ISA) I/O in a computer. Southbridge is one of two chipsets that are collectively called Northbridge /Southbridge. Northbridge controls the processor , memory , Peripheral Component Interconnect (PCI) bus , Level 2 cache , and all Accelerated Graphics Port (AGP) activities. Unlike Northbridge, Southbridge consists of one chip, which sits on Northbridge's PCI bus.

http://www.karbosguide.com/books/pcarchitecture/chapter22.htm
http://www.karbosguide.com/books/pcarchitecture/chapter26.htm

Q.Difference between Southbridge and Northbridge:

Ans.North and south bridge refer to the data channels to the CPU, memory and Hard disk data goes to CPU using the Northbridge. And the mouse, keyboard, CD ROM external data flows to the CPU using the Southbridge.

The Northbridge is the portion of the chipset HUB that connects faster I/O buses (for example, an AGP bus) to the system bus. Northbridge SI also bigger looking then

the Southbridge chip. The Southbridge is the HUB that connects to slower I/O buses (for example, An ISA bus) to the system bus.

The Northbridge and the Southbridge are known as the chipset on the motherboard. These set of chips collectively control the memory cache, external bus, and some peripherals. There is a fast end of the hub, and there is a slow end of the hub. The fast end of the hub is the Northbridge, containing the graphics and memory controller connecting to the system bus. The slower end of the hub is the Southbridge, containing the I/O controller hub.

Note : more point can be added

OverView and Features of PCI , PCI-X and PCI-E AGP

http://en.kioskea.net/contents/403-pci-bus
Q.What is PCI and States its History
Ans. Short for Peripheral Component Interconnect, PCI was introduced by Intel in 1992, revised in 1993 to version 2.0, and later revised in 1995 to PCI 2.1 and is as an expansion to the ISA bus. The PCI bus is a 32-bit (133MBps) computer bus that is also available as a 64-bit bus and was the most commonly found and used computer bus in computers during the late 1990's and early 2000's. Unlike, ISA and earlier expansion cards, PCI follows the PnP specification and therefore does not require any type of jumpers or dip switches. Below is an example of what the PCI slot looks like on a motherboard.

Conventional PCI (**PCI** is from **Peripheral Component Interconnect**, part of the **PCI Local Bus** standard), often shortened to just **PCI**, is a local computer bus for attaching hardware devices in a computer. The PCI bus supports the functions found on a processor bus, but in a standardized format that is independent of any particular processor; devices connected to the PCI bus appear to the processor to be connected directly to the processor bus, and are assigned addresses in the processor's address space

The first version of conventional PCI found in consumer desktop computers was a 32-bit bus operating at 33 Mhz and 5 V signaling, although the PCI 1.0 standard provided for a 64-bit variant as well. Version 2.0 of the PCI standard introduced 3.3 V slots, which are physically distinguished by a flipped physical connector (relative to their 5 V predecessors) to preventing accidental insertion of older cards. Universal cards, which

can operate on both voltages, have two notches. Version 2.1 of the PCI standard introduced optional 66 Mhz operation.

A server-oriented variant of conventional PCI, called **PCI-X** (PCI Extended) operated at higher frequencies, up to 133 Mhz for PCI-X 1.0 and up to 533 Mhz for PCI-X 2.0. An internal connector for laptop cards, called **Mini PCI**, was introduced in version 2.2 of the PCI specification. The PCI bus was also adopted for an external laptop connector standard—the CardBus.The first PCI specification was developed by Intel, but subsequent development of the standard became the responsibility of the PCI Special Interest Group (PCI-SIG).

Conventional PCI and PCI-X are sometimes called **parallel PCI** in order to distinguish them technologically from their more recent successor PCI Express, which adopted a serial, lane-based architecture. Conventional PCI's heyday in the desktop computer market was approximately the decade 1995-2005. PCI and PCI-X have become obsolete for most purposes, however, they are still common on modern desktops for the purposes of backwards compatibility and the low relative cost to produce. Many kinds of devices previously available on PCI expansion cards are now commonly integrated onto motherboards or available serial bus and PCI Express versions.

PCI Express(PCIe) (**Peripheral Component Interconnect Express**), officially abbreviated as **PCIe**, is a high-speed serial computer expansion bus standard designed to replace the older PCI, PCI-X, and AGP bus standards. PCIe has numerous improvements over the aforementioned bus standards, including higher maximum system bus throughput, lower I/O pin count and smaller physical footprint, better performance-scaling for bus devices, a more detailed error detection and reporting mechanism (Advanced Error Reporting (AER)), and native hot-plug functionality. More recent revisions of the PCIe standard support hardware I/O virtualization.

Connector

At least 3 or 4 PCI connectors are generally present on motherboards and can generally be recognised by their standardized white color.

The PCI interface exists in 32 bits with a 124-pin connector, or in 64 bits with a 188-pin connector. There are also two signalling voltage levels:

- 3.3V, for laptop computers
- 5V, for desktop computers

The signalling voltage does not equal the voltage of the motherboard power supply but rather the voltage threshold for the digital encryption of data.

There are 2 types of 32-bit connectors:

- 32-bit PCI connector, 5V:

- 32-bit PCI connector, 3.3V:

The 64-bit PCI connectors offer additional pins and can accommodate 32-bit PCI cards. There are 2 types of 64-bit connectors:

- 64-bit PCI connector, 5V:

- 64-bit PCI connector, 3.3V:

Q.What is ISA?List Feature of ISA Bus?

Ans. The Industry Standard Architecture or ISA bus began as part of IBM's revolutionary PC/XT released in 1981. However, it was officially recognized as "ISA" in 1987 when the IEEE (Institute of Electrical and Electronics Engineers) formally documented standards governing its 16-bit implementation.

This first XT bus was intended to allow the addition of system options which could not be fit onto the motherboard.

This XT bus was completely under the microprocessor's direct control, and its addressing width was limited to the 8-bit level of the processor.

To make the bus useful, control lines were added to signal interrupts for input/output ports. Bus speed was also limited to match the processor. The PC/XT's 8088 was a one-byte wide 4.77 MHz processor. Thus the XT bus, which required two clock cycles for data transfer, was limited to an excruciatingly slow (by today's standards) 2.38 Mbps, that could be curtailed even further if the system was busy with other tasks.

Quatech's first data communication adapters were designed for the PC/ XT, and some of these are being used in older systems running extremely simple, low-speed applications. However, the ISA bus has come a long way since 1981, and its modern incarnation is much better suited to the PCs we use today and the applications we run on them.

Features of ISA

- They are two capabilities that handle data: 8-bit ISA and ISA-16 bits.
- Are physically different expansion slots, the 8 bits is smaller than 16 bits.
- The 16-bit ISA slot also supports 8 bit ISA devices, but not vice versa.
- They have a transfer rate of up to 20 Mbytes / s (MB / s).
- They have a working internal speed of 4.77 MHz, 6 MHz, 8 MHz and 10 MHz
- It has a feature called "bus master" or bus-level control, which allows you to work directly with the RAM.
- ISA could be considered an expansion slot of the second generation.
- This type of expansion slots generate a bottleneck having the higher speed microprocessor.

Q. What is PCI Bus? List its Features?

Ans. The Peripheral Component Interconnect, or PCI Standard, specifies a computer bus for attaching peripheral devices to a computer motherboard.

PCI is short for Peripheral Component Interconnect. The PCI slot is a local system bus standard that was introduced by the Intel Corporation, however, it is not exclusive to any form of processors and PCI slots are found in both Windows PCs and Macs.

PCI slots allow numerous different types of expansion cards to be connected inside a computer to extend the computers functionality. Examples of PCI expansion cards are network cards, graphics cards and sound cards.

The PCI bus is a high speed bus that connects high-performance peripherals like video adapters, disk adapters and network adapters to the chipset. processor and memory.

Unlike previous buses that linked tightly the processor to the expansion bus slot, the PCI bus electronically isolates the processor from the PCI expansion slots. This allows more PCL slots to be supported and removes performance constraints on the use of those slots.

Although the bus speed is slightly slower than PCI Express, the PCI slots are the most common type of slot and found on most motherboards today. If you are installing

a new video card and you are unsure about the slots, stick with the PCI version of this card it will always work.

The most recent motherboards usually provide four or five PCI slots. The chipset provides bridging functions between these 10 buses (The PCI ISA bridge) and between 10 buses and other system buses, including the memory bus. Any system or motherboard today should provide PCI expansion slots in adequate number for s stem needs.

PCI is an evolving standard. Recent models used PCI 2,1 allowing upgradeability. Today's motherboards use PCI 2.2 and PCI 2.3 is already defined and implementations are on their way.

Speed and Width of PCI

- ✓ PCI expansion buses differ in two respects that determine their performance: PCI bus width and bus speed.
- ✓ PCI with 32 bits width at 33.33 MHz generating 133.33 Mbytes/s is found in Desktops and Entry -level servers,
- ✓ PCI with 64 bits width at 66.66 MHz generating 53333 Mbytes/s is more commonly found in Mid-range to High-end Servers.

Feature of PCI

- **Extremely high-speed data transfer:** 32-bit wide data transfer at the rate 33 MI-h gives a maximum throughput of 132 Mega bytes per second. Data transfer at the rate 66 MHz with 64 bit wide data is now being offered.
- **Plug and play facility:** This circumvents the need for an explicit address' for a plug in board. A PCI board inserted in any PCI slot is automatically detected and the required 110 and memory resources are allotted by the system. Thus, there is no risk of clash of resources.
- **New approach:** It moves peripherals off the 1/0 bus and places them closer to the system processor bus, thereby providing faster data transfer between the processor and peripherals.
- **Processor independence:** The PCI local bus fulfills the need for a local bus standard that is not directly dependent on the speed and structure of the processor bus, and that is both reliable and expandable. It is for the first time in PC industry that a common bus, independent of microprocessor and manufacturer, has been adopted.
- **Full multi-master capability :**This allows any PCI master to communicate directly with other PC master/slave.

- **Parity on both data and address lines:** This allows implementation of robust system.
- Support for both SV and 3.3 V operated logic.
- Forward and backward compatibility between 66 MHz and 33MHz PCI.

Q.List the specification and Version of PCI

Ans. Versions of PCI allow (and in the latest versions require) 3.3V slots (keyed differently) on motherboards and allow for cards that are either double keyed for both voltages or even 3.3V only.

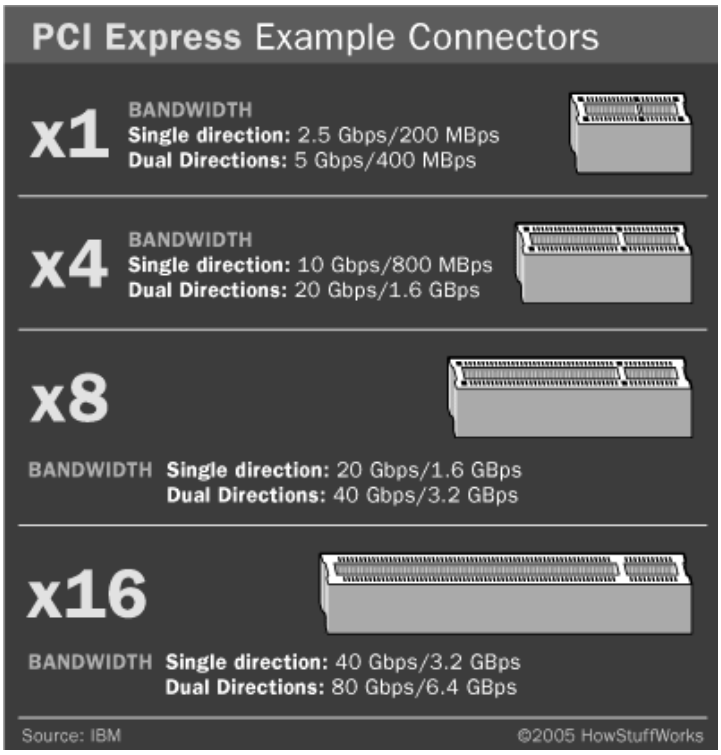
- PC 2.2 allows for 66 MH signaling (requires 3.3 volt signaling) (peak transfer rate of 533 MB/s).
- PCI 2.3 permits use of 3.3 volt and universal keying, but does not allow S volt keyed add in cards.
- PCI 3.0 is the final official standard of the bus, completely removing 5-volt capability.
- PCI-X doubles the width to 64-bit, revises the protocol, and increases the maximum signaling frequency to 133 MHz (peak transfer rate of 1014 MB/s)
- PCI-X 2.0 permits a 266 MHz rate (peak transfer rate of 2035 MB/s) and also 533MHz rate, expands the configuration space to 4096 bytes, adds a 16-bit bus variant and allows for 4.5 volt signaling
- Mini PCI is a new form factor of PCI 2.2 for use mainly inside lap top s
- CardBus is a PC card form factor for 32-bit, 33 MHz PCI
- CompactPCl, uses Eurocard-sized modules plugged into a PCI backplane.
- PC/104-Plus is an industrial bus that uses the PCI signal lines with different connectors.

Specifications of PCI

These specifications represent the most common version of PCI used in normal PCs.

- ✓ 33.33 MHz clock with synchronous transfers.
- ✓ Peak transfer rate of 133 MHz per second for 32-bit bus width (33.33 Mhz x32 bits divide(/) 8 bits/byte =133 MB/s)
- ✓ Peak transfer rate of 266 MB/s for 64-bit bus width

- ✓ 32-bit or 64-bit bus width
- ✓ 32-bit address space (4 gigabytes)
- ✓ 32-bit I/0 port space (now deprecated)
- ✓ 256-byte configuration space
- ✓ 5-volt signaling
- ✓ Reflected-wave switching



Q.state basic difference between PCI , PCI-X and PCI-E Bus

Ans._PCI-X uses a parallel interconnect along a bus that is shared with other PCI-X devices, just like PCI. In fact, PCI-X is best thought of as "**PCI**-e**X** tended", as it is simply an extension of the legacy PCI 32-bit format, with which it is backward-compatible. It differs mainly in the fact that the bus is now 64-bits wide, and runs at higher frequencies (now up to 533MHz, compared to 66MHz - the fastest PCI frequency).

PCI-Express, on the other hand, uses a **serial** interconnect along a switched bus dedicated exclusively to that slot. In this respect, and most others, it uses radically new architecture, having little to do with old PCI. Furthermore, PCI-Express has the unique capability of **multiplying up** individual data "lanes", to produce aggregate interconnects that can deliver up to 16 times the bandwidth of a single lane. This is why you will always see PCI-Express slots referred to as "PCI-Express*4" or "PCI-Express*16" etc.

Q.State Application of PCI bus

Ans. **Applications**

PCI-X has been with us in the server and workstation arena for some time now, as a bus for high-bandwidth server peripherals such as RAID Controllers and Gigabit Ethernet.

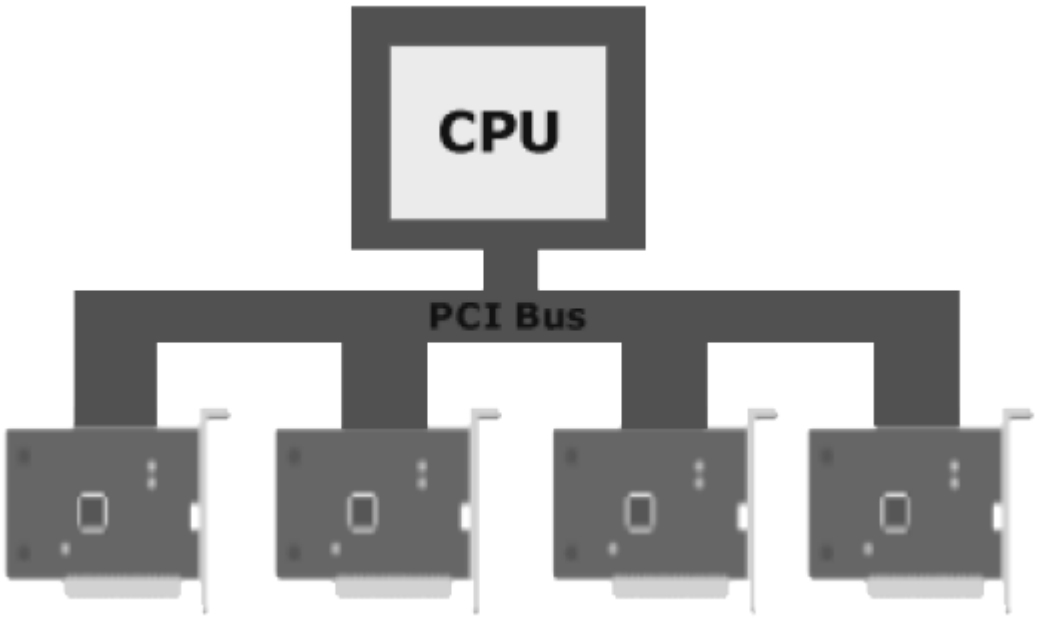
PCI-Express, on the other hand, is brand-new, and is intended to replace AGP in the desktop market and ultimately be the de-facto high-bandwidth peripheral bus across all markets.

Hardware that benefits from 64-bit PCI include:

- **High-performance graphics cards** (PCI-Express only) in the 3D Gaming desktop and graphic intensive workstation markets.
- **U320 SCSI Controllers** for high-speed hard disk access.
- **Multi-port Serial ATA RAID Controllers** for terabyte storage arrays.
- **Gigabit Ethernet** for high-speed networking.
- **IEEE1394b ("Firewire 800")** for ultra-high bandwidth peripherals, such as external hard drives and DV camcorders.

Q.What's wrong with PCI?

Ans.PCI, or Peripheral Component Interconnect was developed by Intel in 1992 and is the local bus used in most PCs until know. PCI uses a shared bus topology to allow for communication among the different devices on the bus i.e. the different PCI devices are attached to the same bus, and share the bandwidth. This diagram explains the situation.



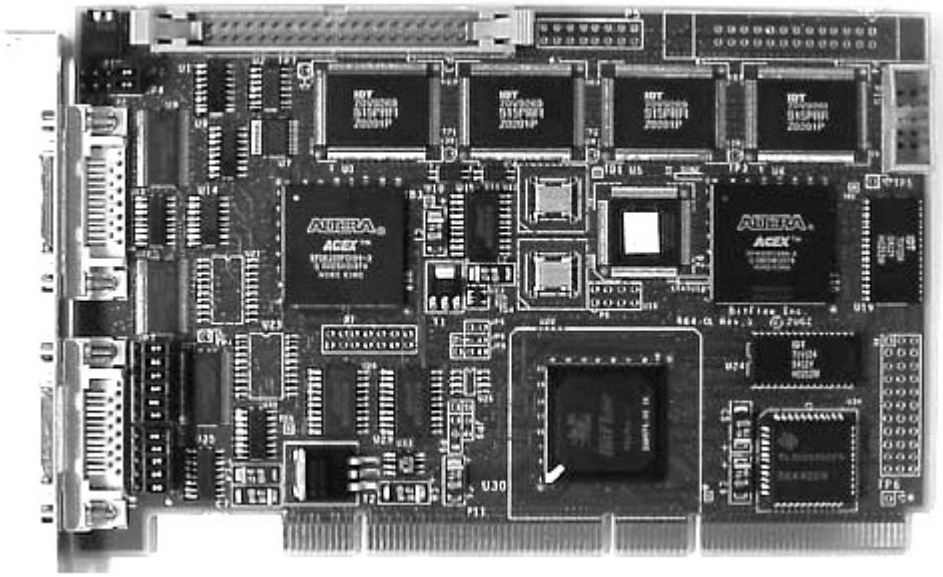
It can run at clock speeds of 33 or 66 MHz. At 32 bits and 33 MHz, it will yield a throughput rate of 133 MBps which is too slow to cater for the latest frame grabbers especially as even this is shared with other PCI devices.

Q.Why is PCI-X an improvement?

Ans.PCI-X stands for PCI Extended.

The PCI-X spec essentially doubled the bus width from 32 bits to 64 bits, thereby increasing bandwidth. The PCI's basic clock rate is increased to 66MHz with a 133MHz variety on the high end, providing another boost to the bandwidth and bringing it up to 1GB/s (at 133MHz).

Having said this PCI-X still suffers from the problem of Shared bus topology and also the faster a bus runs, the more sensitive it becomes to background noise. For this reason manufacturing standards for high-speed buses are exceptionally strict and therefore expensive. The PCI-x slot is physically longer that a PCI Slot.

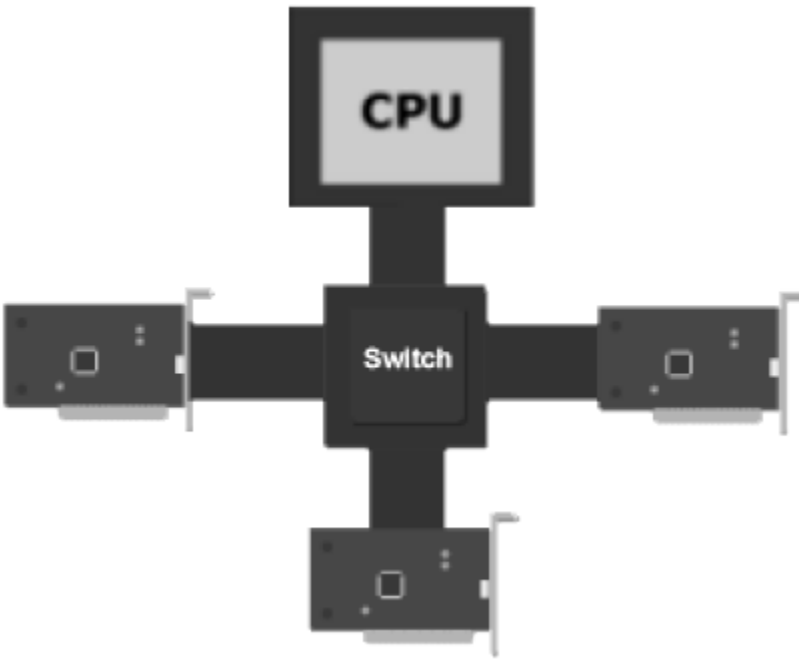


A Bitflow R64 PCI-X frame grabber.

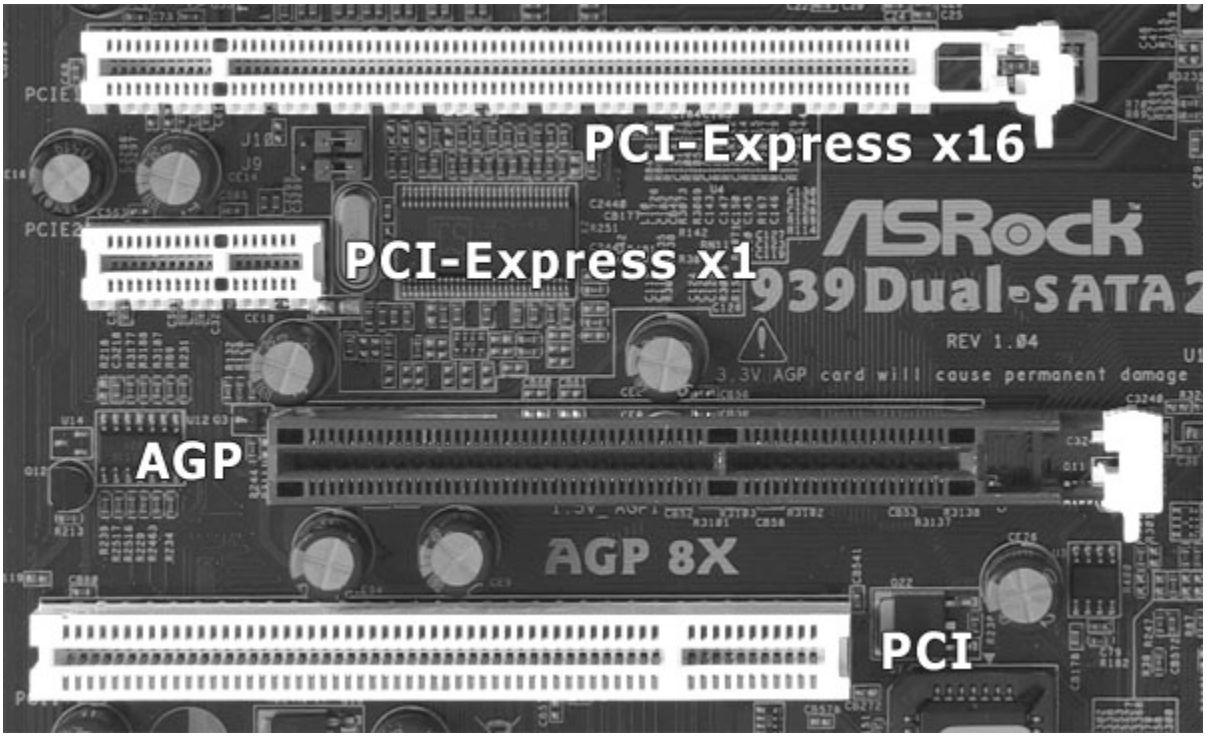
Q.Is PCI-E any better?

Ans.PCI-E stands fro PCI Express and is also known as 3GIO (Third Generation I/O) The most fundamental improvement is the adoption of point-to-point bus topology.

In a point-to-point bus topology, a shared switch replaces the shared bus as the single shared resource by means of which all of the devices communicate. Unlike in a shared bus topology, where the devices must collectively arbitrate among themselves for use of the bus, each device in the system has direct and exclusive access to the switch.



The connections between the devices and the switch is called a link and each link is consists of a number of lanes. Each lane is able to carry data in both directions. The gain in bandwidth is considerable as each lane can carry 2.5Gps in each direction. The PCI Express slot is available in versions of from 1 lane to 32 lanes and are called x1, x2, x4, x8, x16 and x32. The slot and connector are different lengths for each version.



Q.State difference between PCI , PCI-x and PCI-e

Cache Memory

Q.Write Short on Cache Memory

Ans. cachememory is a high-speed memory buffer that temporarily stores data the processor needs, allowing the processor to retrieve that data faster than if it came from main memory. But there is one additional feature of a cache over a simple buffer, and that is intelligence.

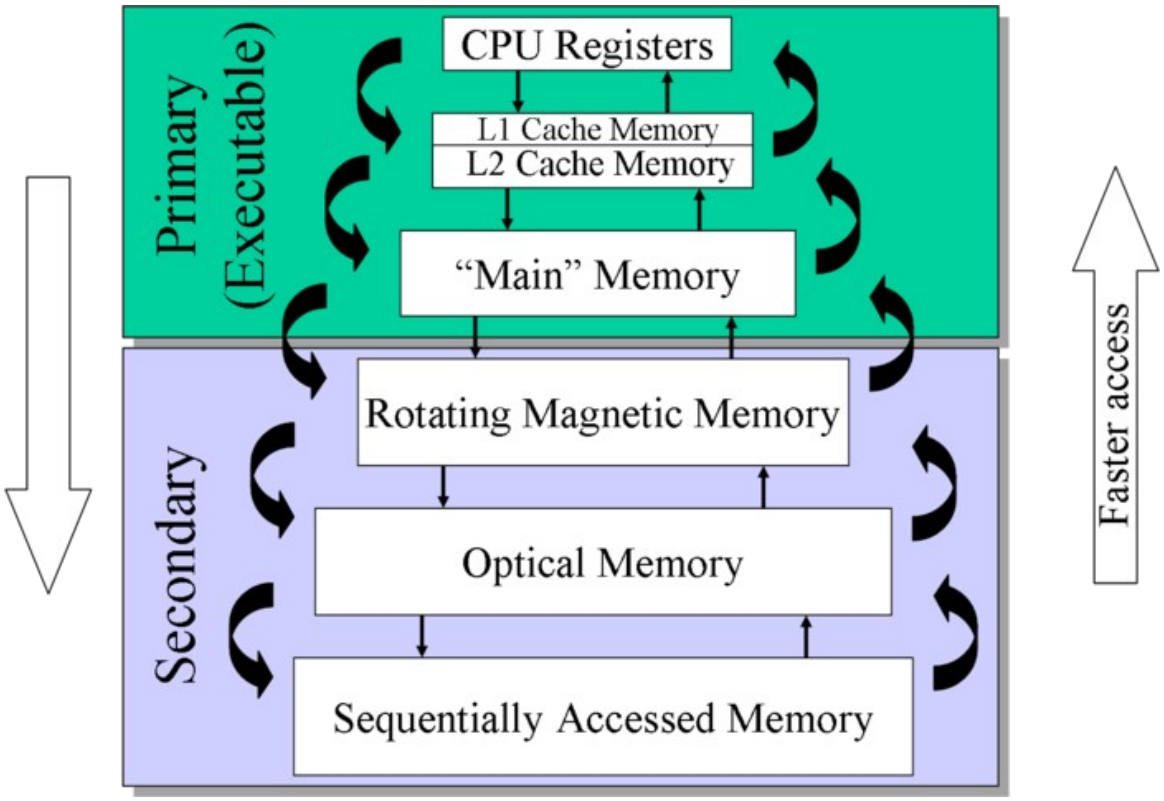
A buffer holds random data, usually on a first-in, first-out basis or a first-in, last-out basis. A cache, on the other hand, holds the data the processor is most likely to need in advance of it actually being needed. This enables the processor to continue working at either full speed or close to it without having to wait for the data to be retrieved from slower main memory. Cache memory is usually made up of static RAM (SRAM) memory integrated into the processor die, although older systems with cache also used chips installed on the motherboard.

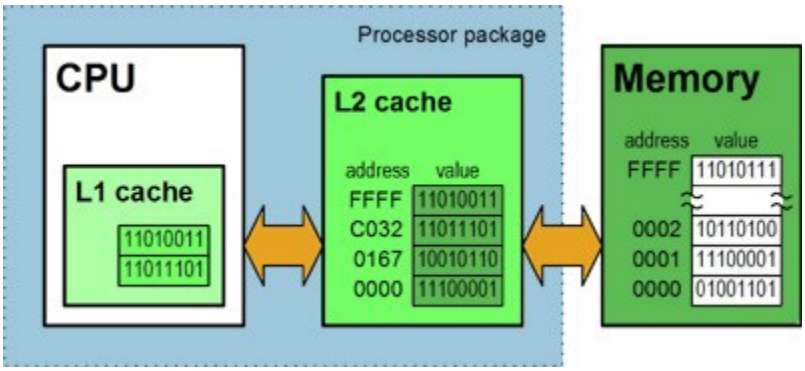
Cache (pronounced cash) memory is extremely fast memory that is built into a computer's central processing unit (CPU), or located next to it on a separate chip. The CPU uses cache memory to store instructions that are repeatedly required to run programs, improving overall system speed. The advantage of cache memory is that the CPU does not have to use the motherboard's system bus for data transfer. Whenever data must be passed through the system bus, the data transfer speed slows to the motherboard's capability. The CPU can process data much faster by avoiding the bottleneck created by the system bus.

As it happens, once most programs are open and running, they use very few resources. When these resources are kept in cache, programs can operate more quickly and efficiently. All else being equal, cache is so effective in system performance that a computer running a fast CPU with little cache can have lower benchmarks than a system running a somewhat slower CPU with more cache. Cache built into the CPU itself is referred to as Level 1 (L1) cache. Cache that resides on a separate chip next to the CPU is called Level 2 (L2) cache. Some CPUs have both L1 and L2 cache built-in and designate the separate cache chip as Level 3 (L3) cache.

Cache that is built into the CPU is faster than separate cache, running at the speed of the microprocessor itself. However, separate cache is still roughly twice as fast as Random Access Memory (RAM). Cache is more expensive than RAM, but it is well worth getting a CPU and motherboard with built-in cache in order to maximize system performance.

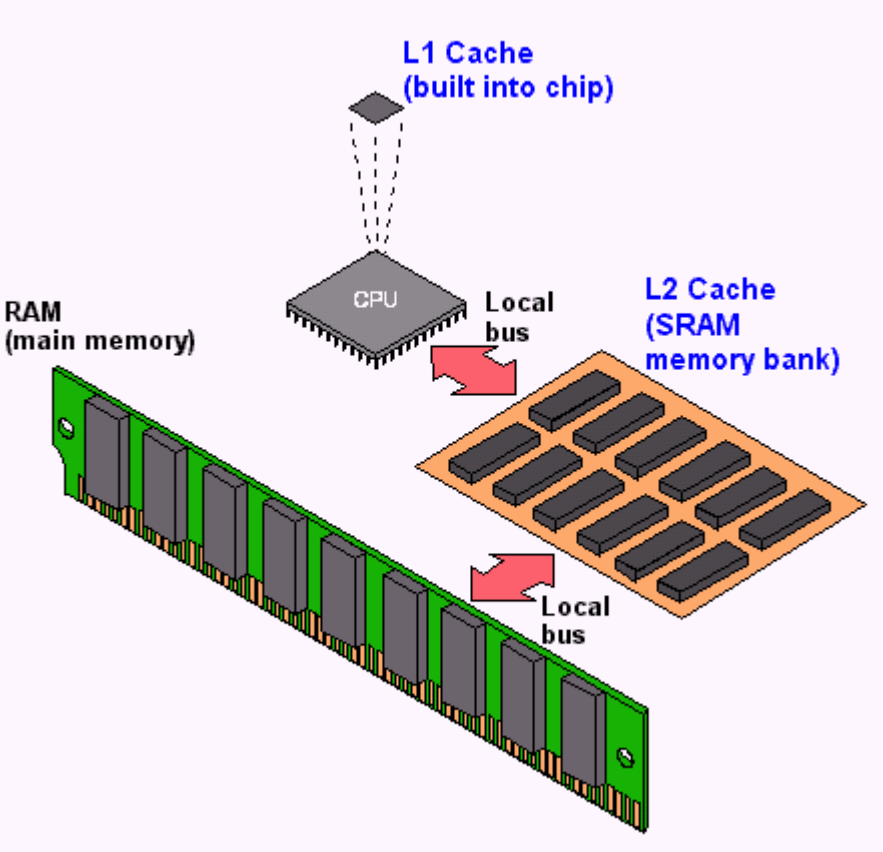
Disk caching applies the same principle to the hard disk that memory caching applies to the CPU. Frequently accessed hard disk data is stored in a separate segment of RAM in order to avoid having to retrieve it from the hard disk over and over. In this case, RAM is faster than the platter technology used in conventional hard disks. This situation will change, however, as hybrid hard disks become ubiquitous. These disks have built-in flash memory caches. Eventually, hard drives will be 100% flash drives, eliminating the need for RAM disk caching, as flash memory is faster than RAM.





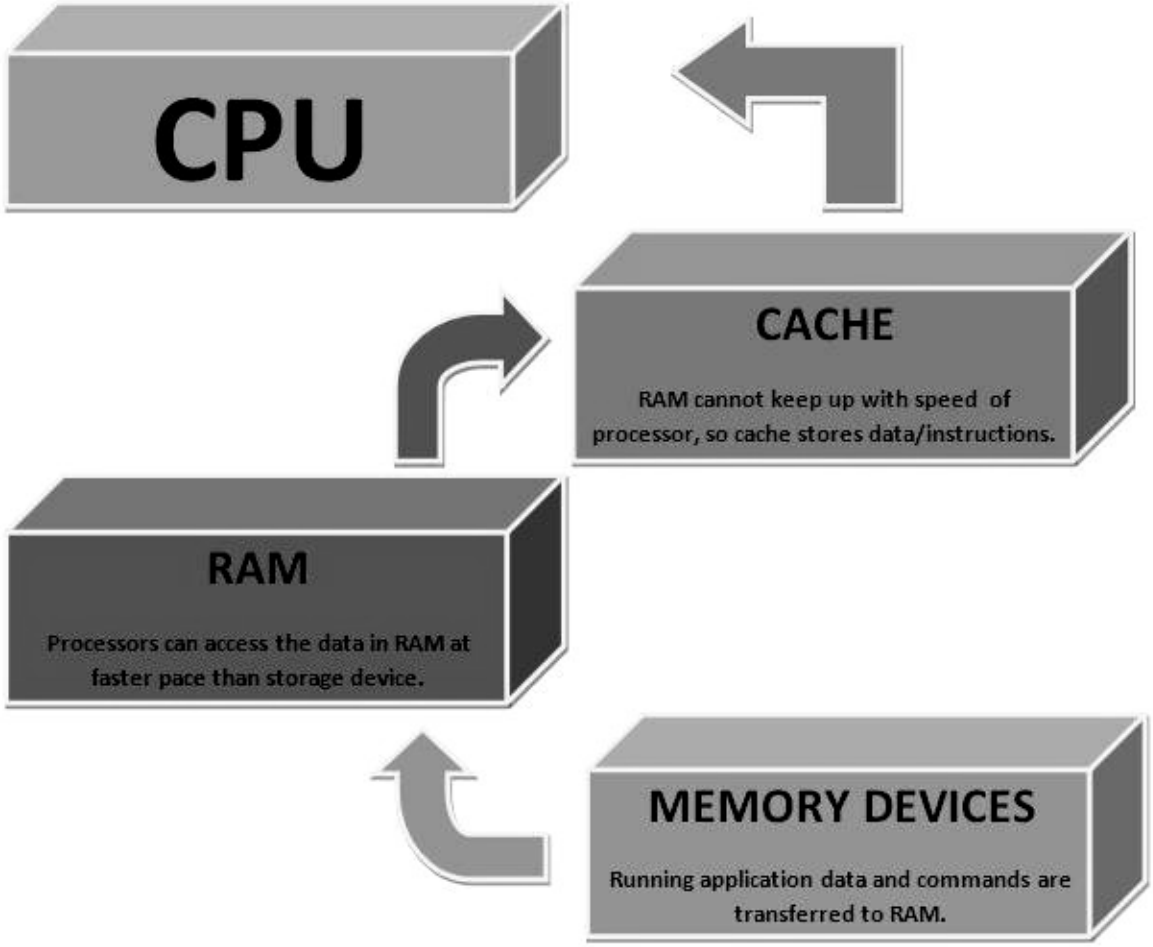
Operation

Let us suppose that the system has cache of three levels (level means that overall cache memory is split into different hardware segments which vary in their processing speed and memory). From RAM data is transferred into cache of 3rd level (L3 cache). L3 cache is a segment of overall cache memory. **L3 cache**is faster than RAM but slower then L2 cache. To further fasten up the process cache of second order L2 cache are used. They are located at immediate vicinity of processor. But in some of the modern processors **L2 cache** is inbuilt making the process faster. It should be noted that it is not necessary that a system has 3 levels of cache; it might have 1 or 2 level of cache. At the core level is cache of first level that is **L1 cache memory**. The commonly used commands/instructions/data is stored in this section of memory. This is built in the processor itself. Thus this is fastest of all the cache memory.

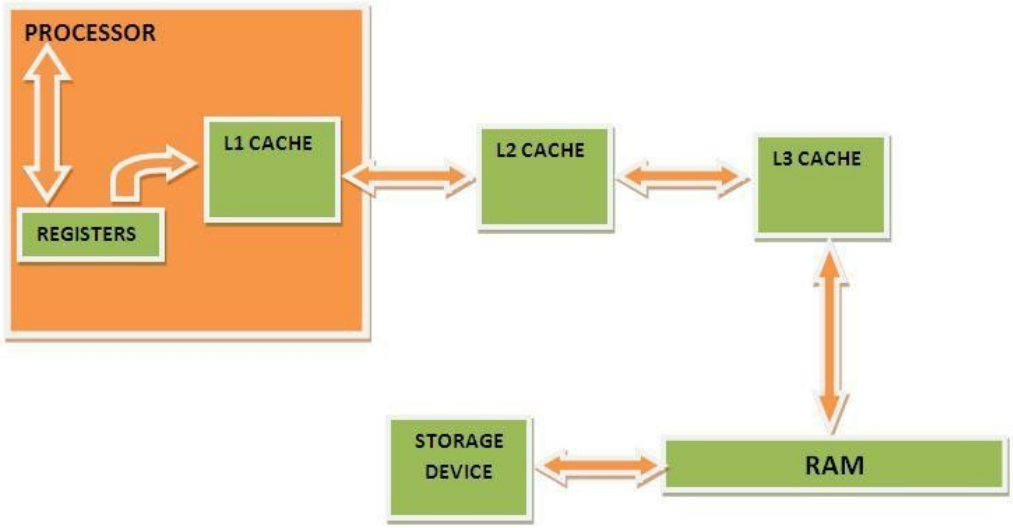


Q.What is Cache State its Purpose?Describe type of cache ? State advantage of cache

Ans.The cache is a smaller, faster memory which stores copies of the data from the most frequently used main memory locations. As long as most memory accesses are to cached memory locations, the avenge latency of memory accesses will be closer to the cache latency than to the latency of main memory.


CPU
CACHE
RAM cannot keep up with speed of processor, so cache stores data/instructions.
RAM
Processors can access the data in RAM at faster pace than storage device.
MEMORY DEVICES
Running application data and commands are transferred to RAM.

When the processor needs to read from or write to a location in main memory, it first checks whether a copy of that data is in the cache. If so, the processor immediately reads from or writes to the cache, which is much faster than reading from or writing to main memory.



The CPU uses cache memory to store instructions that are repeatedly required to run programs, improving overall system speed. The advantage of cache memory is that the CPU does

not have louse the motherboard's system bus for data transfer. Whenever data must be passed through the system bus, the data transfer speed s to the motherboard's capability. The CPU can process data much faster by avoiding the bottleneck created by the system bus.

As it happens, once most programs are open and running, they use very few resources. When these resources are kept in cache, programs can operate more quickly and efficiently. All else being equal, cache is so effective in system performance that a computer running a fast CPU with little cache can have lower benchmarks than a system running a somewhat slower CPU with more cache.

Types of Cache Memory

- Level-1 Cache
- Level-2 Cache
- Level-3 Cache

Level-1

Also called as **L1 cache**, **primary cache**, **internal cache**, or **system cache**. When referring to computer processors, L1 cache is cache that is built into the processor and is the fastest and most expensive cache in the computer. The L1 cache stores the most critical files that need to be executed and is the first thing the processor looks when performing an instruction

Ll, or primary cache, is a small, high-speed cache incorporated right onto the processor's chip. The Li cache typically ranges in size from 8KB to 64KB and uses the high-speed SRAM (static RAM) instead of the slower and cheaper DRAM (dynamic RAM) used for main memory. Using memory cache to hold memory values, or the most recently used data and instructions means the processor can retrieve the data from the cache instead of the system's main memory, which is much slower than the cache memory.

Level 2

L2 is also commonly referred to as **secondary cache** or **external cache**. Unlike Layer 1 cache, L2 cache was located on the motherboard on earlier computers, although with newer processors it is found on the processor chip. When L2 cache is found on the processor, if the cache is also on the motherboard, it is more properly known as L3 cache.

Tip: The L2 cache is on the same processor chip and uses the same die as the CPU, however, it is still not part of the core of the CPU.

L2 cache was first introduced with the Intel Pentium and Pentium Pro computers and has been included with ever process since, with the exception of the early versions of Celeron processor. This cache is not as fast as the L1 cache, but is only slightly slower since it is still located on the same processor chip, and is still faster than the computers memory. The L2 cache is the second thing the computer looks at when performing instructions.

L2,. or secondary cache, is memory between the RAM and the CPU (but not on the CPU chip itself and is bigger than the primary cache (typically 64KB to 2MB). L2 ATC (Advanced Transfer Cache) uses micro-architectural improvements, which provide a higher data bandwidth interface between the L2 cache and the processor core, and is completely scaleable with the processor core frequency. The L2 cache is also a unified, non-blocking cache, which improves performance over cache-on-motherboard solutions through a dedicated 64-bit
cache

Level-3

L3 Cache is Cache found on the motherboard instead of the processor on earlier computers. With today's computers this type of cache is a cache that is found on the same chip and die as the processors. In the below picture of the Intel Core i7-3960X Processor die, is an example of a processor chip containing six cores (CPUs) and the shared L3 Cache. As can be seen in the picture, the L3 cache is shared between all cores (CPUs) and is very large in comparison to what an L1 or L2 cache would be on the same chip because it is cheaper although slower.

Since more manufacturers are beginning to include L2 cache into their architectures, L3 cache is slowly replacing the L2 cache function the extra cache built into the motherboards between the CPU and the main memory (old L2 cache definition) is now being called the L3 cache.

Some manufacturers have proprietary L3 cache designs already, but most desktop and notebook computers do not offer this feature yet. Micron has developed a chip set with 8MB of on-chip DRAM in the north bridge chip that acts as an L3 cache, but offering an L3 cache as standard equipment is still a future prospect.

Advantage of Cache

- The cache memory enhances the speed of system or improving performance.
- Cache memory reduces a traditional system bottleneck.
- As the cache memory lies on the same chip (For LI cache) the access time is very small.

- The same block of data which are stored on the main memory resides on the cache. Thus the instructions takes less time to execute.
- The CPU and the cache are connected with a local bus which is of high capacity and speed due to which the data transfer is quick.
- Cache memory is intelligent memory.
- It holds current working set of code and data.
- It reduces wait state or no wait states (LI cache) in system.

* The initial level of storage on a processor are the registers. The registers are where the actually processing input and output takes place.

* **L1 cache** – Then the level 1 cache comes next. It is logically the closest high speed memory to the CPU core / registers. It usually runs at the full speed (meaning the same as the CPU core clockspeed). L1 often comes in size of 8kB, 16kB, 32kB, 64kB or 128kB. But, it is very high speed even though the amount is relatively small.

* **L2 cache** – The next level of cache is L2, or level 2. Nowadays L2 is larger than L1 and it often comes in 256kB, 512kB and 1,024MB amounts. L2 often runs at 1/4, 1/2 or full speed in relation to the CPU core clockspeed.

* **L3 cache** – Level 3 cache is something of a luxury item. Often only high end workstations and servers need L3 cache. Currently for consumers only the Pentium 4 Extreme Edition even features L3 cache. L3 has been both "on-die", meaning part of the CPU or "external" meaning mounted near the CPU on the motherboard. It comes in many sizes and speeds.

Q.Describe different types of Cache Memories with respect to processor
Ans. There are Three Types

1.Level-1
2.Level-2
3.Level-3

Internal Level 1 Cache

All modern processors starting with the 486 family include an integrated L1 cache and controller. The integrated L1 cache size varies from processor to processor, starting at 8 KB for the original 486DX and now up to 128 KB or more in the latest processors.

Note : Multi-core processors include separate L1 caches for each processor core. Also, L1 cache is divided into equal amounts for instructions and data.

To understand the importance of cache, you need to know the relative speeds of processors and memory. The problem with this is that processor speed usually is expressed in MHz or GHz (millions or billions of cycles per second), whereas memory speeds are often expressed in nanoseconds (billionths of a second per cycle). Most newer types of memory express the speed in either MHz or in megabyte per second (MB/s) bandwidth (throughput).

Both are really time- or frequency-based measurements. You will note that a 233 MHz processor equates to 4.3-nanosecond cycling, which means you would need 4 ns memory to keep pace with a 200 MHz CPU. Also, note that the motherboard of a 233 MHz system typically runs at 66 MHz, which corresponds to a speed of 15 ns per cycle and requires 15 ns memory to keep pace. Finally, note that 60 ns main memory (common on many Pentium-class systems) equates to a clock speed of approximately 16 MHz. So, a typical Pentium 233 system has a processor running at 233 MHz (4.3 ns per cycle), a motherboard running at 66 MHz (15 ns per cycle), and main memory running at 16 MHz (60 ns per cycle). This might seem like a rather dated example, but in a moment, you will see that the figures listed here make it easy for me to explain howcache memory works.

Because L1 cache is always built into the processor die, it runs at the full-core speed of the processor internally. By full-core speed, I mean this cache runs at the higher clock multiplied internal processor speed rather than the external motherboard speed. This cache basically is an area of fast memory built into the processor that holds some of the current working set of code and data. Cache memory can be accessed with no wait states because it is running at the same speed as the processor core.

Cache is even more important in modern processors because it is often the only memory in the entire system that can truly keep up with the chip. Most modern processors are clock multiplied, which means they are running at a speed that is really a multiple of themotherboard into which they are plugged. The only types of memory matching the full speed of the processor are the L1, L2, and L3 caches built into the processor core.

If the data that the processor wants is already in L1 cache, the CPU does not have to wait. If the data is not in the cache, the CPU must fetch it from the Level 2 or Level 3 cache or (in less sophisticated system designs) from the system bus—meaning main memory directly.

According to Intel, the L1 cache in most of its processors has approximately a 90% hit ratio. (Some processors, such as the Pentium 4, are slightly higher.) This means

that the cache has the correct data 90% of the time, and consequently the processor runs at full speed (233 MHz in this example) 90% of the time. However, 10% of the time the cache controller guesses incorrectly, and the data has to be retrieved out of the significantly slower main memory, meaning the processor has to wait. This essentially throttles the system back to RAM speed, which in this example was 60 ns or 16 MHz.

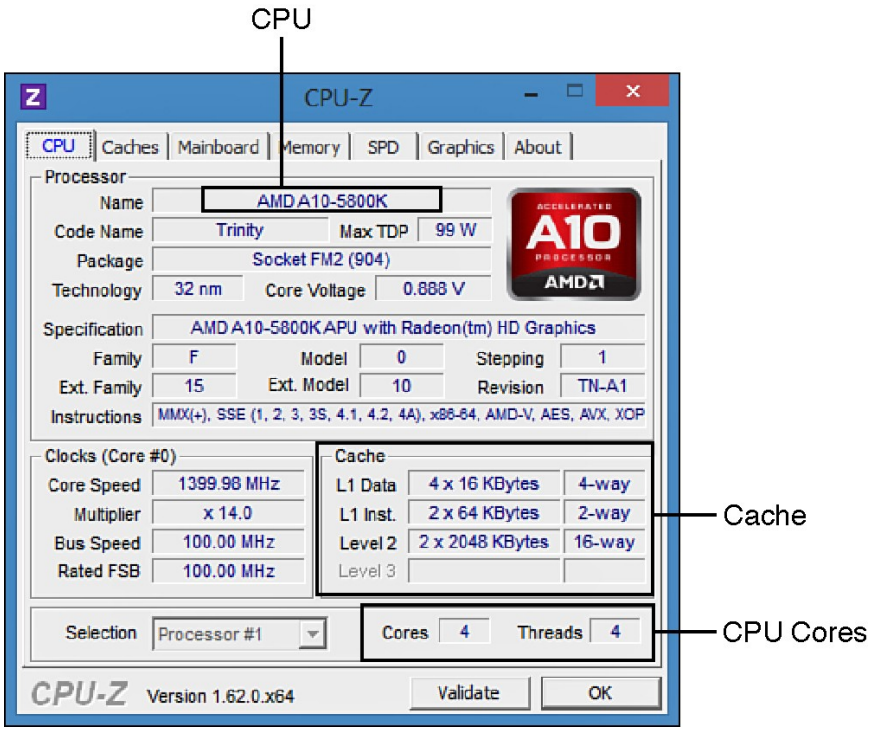
In this analogy, the processor was 14 times faster than the main memory. Memory speeds have increased from 16 MHz (60 ns) to 333 MHz (3.0 ns) or faster in the latest systems, but processor speeds have also risen to 3 GHz and beyond. So even in the latest systems, memory is still 7.5 or more times *slower* than the processor. Cache is what makes up the difference.

The main feature of L1 cache is that it has always been integrated into the processor core, where it runs at the same speed as the core. This, combined with the hit ratio of 90% or greater, makes L1 cache important for system performance.

Level 2 Cache

In an actual Pentium class (Socket_7) system, the L2 cache is mounted on the motherboard, which means it runs at motherboard speed (66 MHz, or 15 ns in this example).

All modern processors have integrated L2 cache that runs at the same speed as the processor core, which is also the same speed as the L1 cache. The screenshot below illustrates the cache types and sizes in the AMD A10-5800Kprocessor, as reported by CPU-Z.



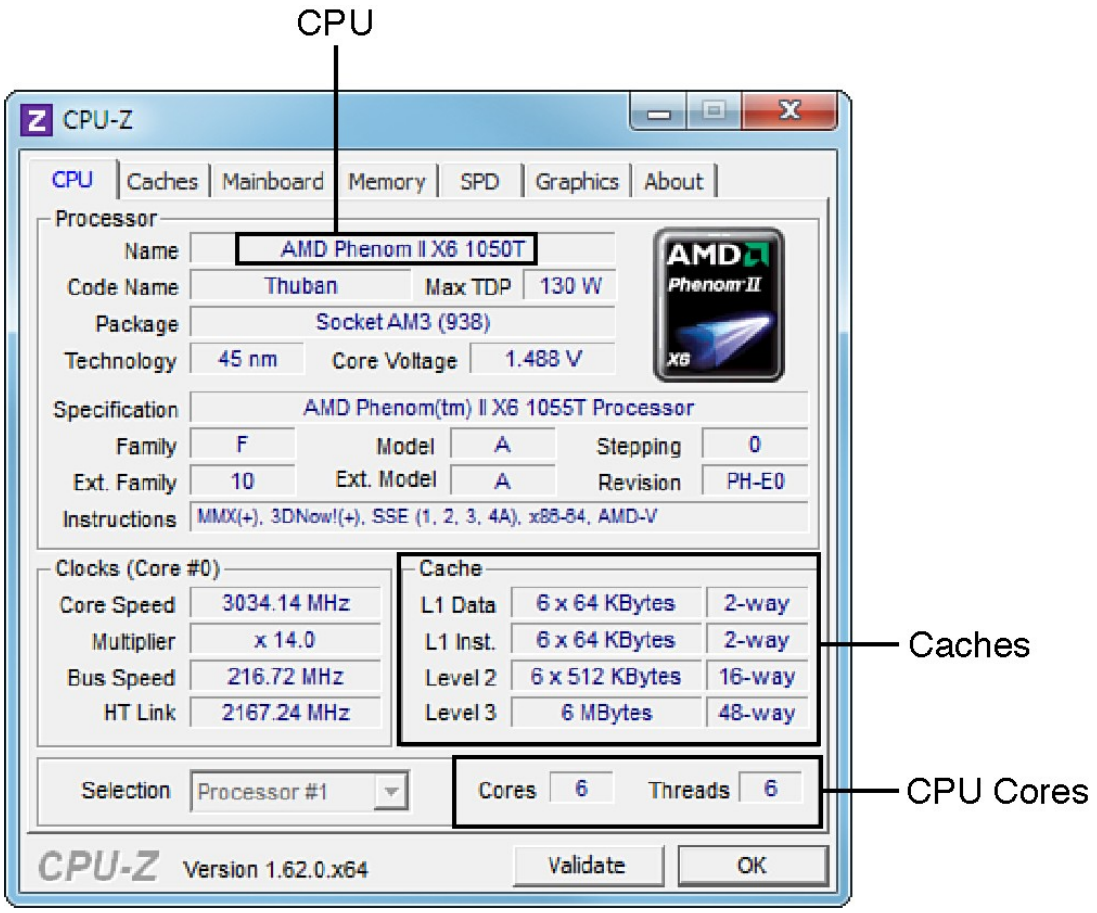
The AMD A10-5800K processor is a quad-core processor with L1 and L2 cache.

Level 3 Cache

Most late-model mid-range and high-performance processors also contain a third level of cache known as *L3 cache*. In the past, relatively few processors had L3 cache, but it is becoming more and more common in newer and

faster multicore processors such as the Intel Core i7 and AMD Phenom II and FX processors.

L3 cache proves especially useful in multicore processors, where the L3 is generally shared among all the cores. Both Intel and AMD use L3 cache in most of their current processors because of the benefits to multicore designs.



+Cache Information for the Intel Core i5-2500 (Sandy Bridge)

These screenshots illustrate two examples of six-core processors with L1, L2, and L3 cache from both Intel (above) and AMD (below):

Cache information for the AMDPhenom II X6 1055T

Just as with the L1 cache, most L2 caches have a hit ratio also in the 90% range; therefore, if you look at the system as a whole, 90% of the time it runs at full speed (233 MHz in this example) by retrieving data out of the L1 cache. Ten percent of the time it slows down to retrieve the data from the L2 cache. Ninety percent of the time the processor goes to the L2 cache, the data is in the L2, and 10% of that time it has to go to the slow main memory to get the data because of an L2 cache miss. So, by combining both caches, our sample system runs at full processor speed 90% of the time (233 MHz in this case), at motherboard speed 9% (90% of 10%) of the time (66 MHz in this case),

and at RAM speed about 1% (10% of 10%) of the time (16 MHz in this case). You can clearly see the importance of both the L1 and L2 caches; without them the system uses main memory more often, which is significantly slower than the processor.

Important Point to be noted.

Spending money doubling the performance of either the main memory (RAM) or the L2 cache, which would you improve? Considering that main memory is used directly only about 1% of the time, if you doubled performance there, you would double the speed of your system only 1% of the time! That doesn't sound like enough of an improvement to justify much expense. On the other hand, if you doubled L2 cache performance, you would be doubling system performance 9% of the time, which is a much greater improvement overall. I'd much rather improve L2 than RAM performance. The same argument holds true for adding and increasing the size of L3 cache, as many recent processors from AMD and Intel have done.

The processor and system designers at Intel and AMD know this and have devised methods of improving the performance of L2 cache. In Pentium (P5) class systems, the L2 cache usually was found on the motherboard and had to run at motherboard speed. Intel made the first dramatic improvement by migrating the L2 cache from the motherboard directly into the processor and initially running it at the same speed as the main processor. The cache chips were made by Intel and mounted next to the main processor die in a single chip housing. This proved too expensive, so with the PentiumII, Intel began using cache chips from third-party suppliers such as Sony, Toshiba, NEC, and Samsung. Because these were supplied as complete packaged chips and not raw die, Intel mounted them on a circuit board alongside the processor. This is why the Pentium II was designed as a cartridge rather than what looked like a chip.

Q.Comparision of Cache Memories

OverView and Features of SDRAM , DDR, DDR2 and DDR3

Q.What is SDRAM

Ans. Synchronous dynamic random access memory (SDRAM) is dynamic random access memory (DRAM) with an interface synchronous with the system bus carrying data between the CPU and the memory controller hub. SDRAM has a rapidly responding synchronous interface, which is in sync with the system bus. SDRAM waits for the clock signal before it responds to control inputs.

SDRAM preceded double data rate (DDR). The newer interface of DRAM has a double data transfer rate using both the falling and rising edges of the clock signal. This is called dual-pumped, double pumped or double transition. There are three significant characteristics differentiating SDRAM and DDR:

1. The main difference is the amount of data transmitted with each cycle, not the speed.
2. SDRAM sends signals once per clock cycle. DDR transfers data twice per clock cycle. (Both SDRAM and DDR use the same frequencies.)
3. SDRAM uses one edge of the clock. DDR uses both edges of the clock.

SDRAM has a 64-bit module with long 168-pin dual inline memory modules (DIMMs). SDRAM access time is 6 to 12 nanoseconds (ns). SDRAM is the replacement for dynamic random access memory (DRAM) and EDO RAM. DRAM is a type of random access memory (RAM) having each bit of data in an isolated component within an integrated circuit. Older EDO RAM performed at 66 MHz.

Detailed Explanation

With older clocked electronic circuits, the transfer rate was one per full cycle of the clock signal. This cycle is called rise and fall. A clock signal changes two times per transfer, but the data lines change no more than one time per transfer. This restriction can cause integrity (data corruption and errors during transmission) when high bandwidths are used. SDRAM transmits signals once per clock cycle. The newer DDR transmits twice per clock cycle.

SDRAM is improved DRAM with a synchronous interface waiting for a clock pulse before it responds to data input. SDRAM uses a feature called pipelining, which accepts new data before finishing processing previous data. A delay in data processing is called latency.

DRAM technology has been used since the 1970's. In 1993, SDRAM was implemented by Samsung with model KM48SL2000 synchronous DRAM. By 2000, DRAM was replaced by SDRAM. In the beginning SDRAM was slower than burst EDO DRAM

because of the extra logic features. But the benefits of SDRAM allowed more than one set of memory, which increased the bandwidth efficiency.

With the introduction of DDR, SDRAM quickly began to fade out of use because DDR was cheaper and more cost effective. The SDRAM used a 168-pin while the DDR module used a 184-pin. SDRAM modules used a voltage of 3.3V and DDR used 2.6V, producing less heat.

Features of SDRAM

1) Speed is between 100Mhz to 133Mhz
2) Sizes Available is 32Mb,64Mb,128Mb,256Mb,512Mb,1Gb
3) Operating Voltage is 3.3V
4) Synchronous Architecture
5) SDRAM is synchronized with the computer's clock to allow it to send instructions more efficiently by joining a pipeline of other instructions the computer is processing.
6) The pipelining of information in a computer allows it to receive another command before it has finished processing the previous command.
7) This allows SDRAM to operate at much higher speeds, making it the most popular form of RAM offered on computers.
8) SDRAM module are of 168 pin
9) It prefetch 1 bit at a time

DDR (Dual Data Rate)

Q.What is DDR

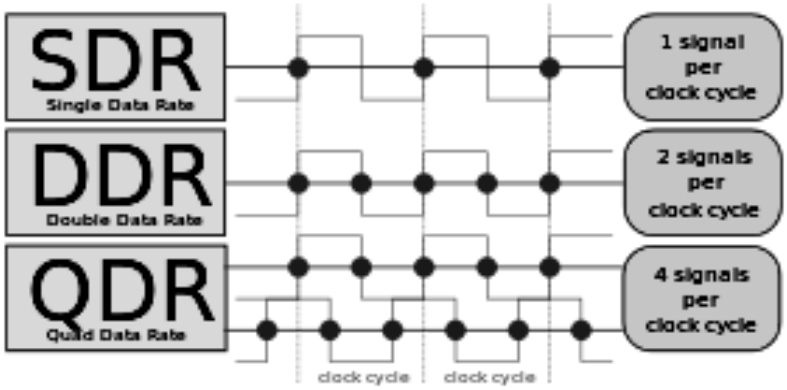
Ans. In computing, a computer bus operating with double data rate (DDR) transfers data on both the rising and falling edges of the clock signal.This is also known as double pumped, dual-pumped, and double transition. The term toggle mode is used in the context of NAND flash memory.

The simplest way to design a clocked electronic circuit is to make it perform one transfer per full cycle (rise and fall) of a clock signal. This, however, requires that the

clock signal changes twice per transfer, while the data lines change at most once per transfer. When operating at a high bandwidth, signal integrity limitations constrain the clock frequency. By using both edges of the clock, the data signals operate with the same limiting frequency, thereby doubling the data transmission rate.

This technique has been used for microprocessor front side busses, Ultra-3 SCSI, graphics RAM (the AGP bus and GDDR), main memory (both RDRAM and DDR1 through DDR4), and the HyperTransport bus on AMD's Athlon 64 processors. It is more recently being used for other systems with high data transfer speed requirements – as an example, for the output of analog-to-digital converters (ADCs).

DDR should not be confused with dual channel, in which each memory channel accesses two RAM modules simultaneously. The two technologies are independent of each other and many motherboards use both, by using DDR memory in a dual channel configuration.



The 184-pin DDR RAM dual in-line memory modules (DIMMS) only work properly in a motherboard designed for their use. While this RAM comes in various speeds, installing a version faster than a motherboard can support is a waste of money, since it will only run as fast as the motherboard permits. It is visually differentiated from SDRAM in that SDRAM is a 168-pin DIMM with a double notch at the bottom along the pins — one notch just off-center, the other offside. The 184-pin DDR SDRAM has a single off-center notch.

DDR RAM is generally made for processors 1GHz and faster. Designations like *PC1600 DDR SDRAM* and *PC2100 DDR SDRAM* coincide with particular FSB and CPU speeds. RAM manufacturers use different schemes to designate processor speed, and the various technicalities in RAM designations and standards can be confusing. Computer users should check their motherboard manual to see what RAM type is compatible with their system before purchasing memory.

SDR SDRAM (Single Data Rate synchronous DRAM)

This type of SDRAM is slower than the DDR variants, because only one word of data is transmitted per clock cycle (single data rate). But this type is also faster than its predecessors EDO-RAM and FPM-RAM which took typically 2 or 3 clocks to transfer one word of data.

DDR-1 SDRAM

While the access latency of DRAM is fundamentally limited by the DRAM array, DRAM has very high potential bandwidth because each internal read is actually a row of many thousands of bits. To make more of this bandwidth available to users, a double data rate interface was developed. This uses the same commands, accepted once per cycle, but reads or writes two words of data per clock cycle. The DDR interface accomplishes this by reading and writing data on both the rising and falling edges of the clock signal. In addition, some minor changes to the SDR interface timing were made in hindsight, and the supply voltage was reduced from 3.3 to 2.5 V. As a result, DDR SDRAM is not backwards compatible with SDR SDRAM.

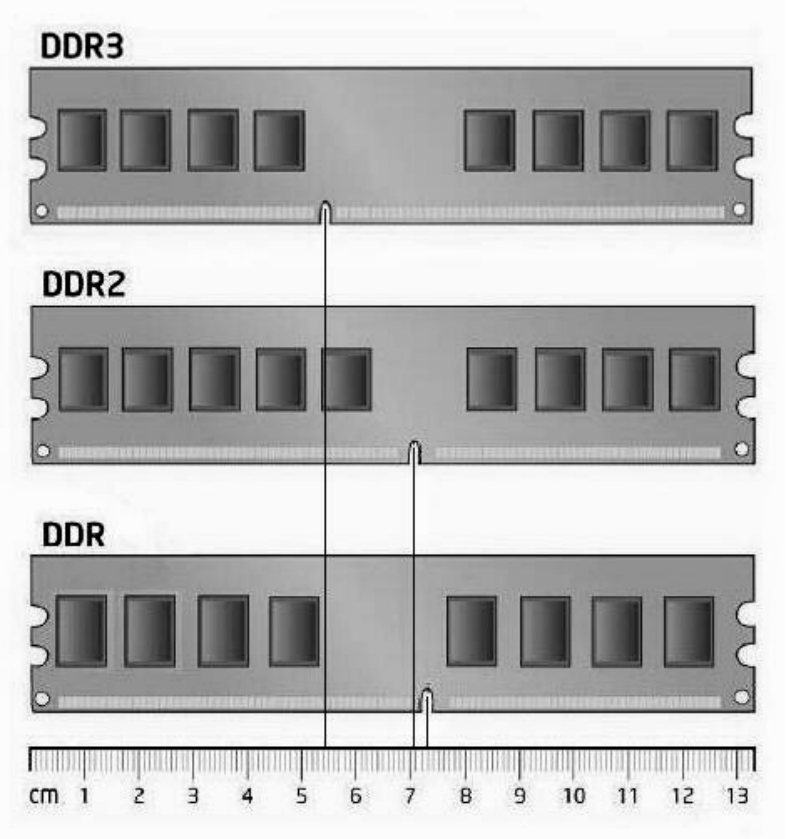
DDR SDRAM (sometimes called DDR1 for greater clarity) doubles the minimum read or write unit; every access refers to at least two consecutive words.

Typical DDR SDRAM clock rates are 133, 166 and 200 MHz (7.5, 6, and 5 ns/cycle), generally described as DDR-266, DDR-333 and DDR-400 (3.75, 3, and 2.5 ns per bit). Corresponding 184-pin DIMMs are known as PC-2100, PC-2700 and PC-3200.

Features of DDR-1

i. **184 pin module**
ii. **Speed – 100Mhz , 133 Mhz , 166Mhz , 200Mhz**
iii. **Operating voltage 2.5v**
iv. **Synchronous Architecture**
v. **It prefetches 2 bits at a time**
vi. **Its is Double faster then SDRAM**

DDR2 SDRAM



DDR2 SDRAM is very similar to DDR SDRAM, but doubles the minimum read or write unit again, to 4 consecutive words. The bus protocol was also simplified to allow higher performance operation. (In particular, the "burst terminate" command is deleted.) This allows the bus rate of the SDRAM to be doubled without increasing the clock rate of internal RAM operations; instead, internal operations are performed in units 4 times as wide as SDRAM. Also, an extra bank address pin (BA2) was added to allow 8 banks on large RAM chips.

Typical DDR2 SDRAM clock rates are 200, 266, 333 or 400 MHz (periods of 5, 3.75, 3 and 2.5 ns), generally described as DDR2-400, DDR2-533, DDR2-667 and DDR2-800 (periods of 2.5, 1.875, 1.5 and 1.25 ns). Corresponding 240-pin DIMMS are known as PC2-3200 through PC2-6400. DDR2 SDRAM is now available at a clock rate of 533 MHz generally described as DDR2-1066 and the corresponding DIMMs are known as PC2-8500 (also named PC2-8600 depending on the manufacturer).

Note that because internal operations are at 1/2 the clock rate, DDR2-400 memory (internal clock rate 100 MHz) has somewhat higher latency than DDR-400 (internal clock rate 200 MHz).

Features of DDR-2

i. **It is 240 pin module**
ii. **Speed – 400Mhz , 533Mhz,667Mhz and above**
iii. **High Bandwidth**
iv. **Synchronous Architecture**
v. **Operating at 1.8V**
vi. **Prefetch 4bit at a time**

DDR3 SDRAM

DDR3 continues the trend, doubling the minimum read or write unit to 8 consecutive words. This allows another doubling of bandwidth and external bus rate without having to change the clock rate of internal operations, just the width. To maintain 800–1600 M transfers/s (both edges of a 400–800 MHz clock), the internal RAM array has to perform 100–200 M fetches per second.

Again, with every doubling, the downside is the increased latency. As with all DDR SDRAM generations, commands are still restricted to one clock edge and command latencies are given in terms of clock cycles, which are half the speed of the usually quoted transfer rate (a CAS latency of 8 with DDR3-800 is 8/(400 MHz) = 20 ns, exactly the same latency of CAS2 on PC100 SDR SDRAM).

DDR3 memory chips are being made commercially, and computer systems using them were available from the second half of 2007, with significant usage from 2008 onwards. Initial clock rates were 400 and 533 MHz, which are described as DDR3-800 and DDR3-1066 (PC3-6400 and PC3-8500 modules), but 667 and 800 MHz, described as DDR3-1333 and DDR3-1600 (PC3-10600 and PC3-12800 modules) are now common

Features of DDR-3

i. **New Pin ie asynchronous RESET pin was introduced**
ii. **On-DIMM mirror friendly DRAM pin out**
iii. **Read and Write calibration**
iv. **Operating voltage 1.5V**
v. **Prefetch 8bit at a time**
vi. **Speed- 800Mhz , 1066Mhz , 1333Mhz , 1600Mhx and Above**

DDR4 SDRAM

DDR4 SDRAM is the successor to DDR3 SDRAM. It was revealed at the Intel Developer Forum in San Francisco in 2008, and is due to be released to market during 2011. The timing has varied considerably during its development - it was originally expected to be released in 2012, and later (during 2010) expected to be released in 2015, before samples were announced in early 2011 and manufacturers began to announce that commercial production and release to market was anticipated in 2012. DDR4 is expected to reach mass market adoption around 2015, which is comparable with the approximately 5 years taken for DDR3 to achieve mass market transition over DDR2.

The new chips are expected to run at 1.2 V or less, versus the 1.5 V of DDR3 chips, and have in excess of 2 billion data transfers per second. They are expected to be introduced at frequency rates of 2133 MHz, estimated to rise to a potential 4266 MHz and lowered voltage of 1.05 V by 2013.

DDR4 will not double the internal prefetch width again, but will use the same 8n prefetch as DDR3. Thus, it will be necessary to interleave reads from several banks to keep the data bus busy.

In February 2009, Samsung validated 40 nm DRAM chips, considered a "significant step" towards DDR4 development since as of 2009, current DRAM chips were only beginning to migrate to a 50 nm process. In January 2011, Samsung announced the completion and release for testing of a 30 nm 2 GB DDR4 DRAM module. It has a maximum bandwidth of 2.13 Gbit/s at 1.2 V, uses pseudo open drain technology and draws 40% less power than an equivalent DDR3 module.

Feature of SDRAM , DDR and it versions0

Type	Feature changes
SDRAM	Vcc = 3.3 V Signal: LVTTL Access is ≥2 words Double clocked
DDR1	Vcc = 2.5 V 2.5 - 7.5 ns per cycle Signal: SSTL_2 (2.5V)[18] Access is ≥4 words "Burst terminate" removed
DDR2	4 units used in parallel 1.25 - 5 ns per cycle Internal operations are at 1/2 the clock rate. Signal: SSTL_18 (1.8V)[18] Access is ≥8 words
DDR3	Signal: SSTL_15 (1.5V)[18] Much longer CAS latencies
DDR4	Vcc ≤ 1.2 V point-to-point (single module per channel)

Stub Series Terminated Logic (SSTL) is a group of electrical standards for driving transmission lines commonly used with DRAM based DDR memory IC's and memory modules. SSTL is primarily designed for driving the DDR (double-data-rate) SDRAM modules used in computer memory

Comparision of DDR-1 DDr-2 and DDR-3

Items	DDR3 SDRAM	DDR2 SDRAM	DDR SDRAM
Clock frequency	400/533/667/ 800MHz	200/266/333/400MHz	100/133/166/200MHz
Transfer data rate	800/1066/1333/ 1600Mbps	400/533/667/800Mbps	200/266/333/400Mbps
I/O width	x4/x8/x16	x4/x8/x16	x4/x8/x16/x32
Prefetch bit width	8-bits	4-bits	2-bits
Clock input	Differential clock	Differential clock	Differential clock
Burst length	8, 4 (Burst chop)	4, 8	2, 4, 8
Data strobe	Differential data strobe	Differential data strobe	Single data strobe
Supply voltage	1.5V	1.8V	2.5V
Interface	SSTL_15	SSTL_18	SSTL_2
/CAS latency (CL)	5, 6, 7, 8, 9, 10 clock	3, 4, 5 clock	2, 2.5, 3 clock
On die termination (ODT)	Supported	Supported	Unsupported
Component package	FBGA	FBGA	TSOP(II) / FBGA / LQFP

	SDRAM	DDR1	DDR2	DDR3
Power supply voltage	3.3V	VDDQ = 2.5V VTT = 1/2 x VDDQ VREF = 1/2 x VDDQ	VDDQ = 1.8V VTT = 1/2 x VDDQ VREF = 1/2 x VDDQ	VDDQ = 1.5V VTT = 1/2 x VDDQ VREF = 1/2 x VDDQ
Interface	LVTTL	SSTL_2	SSTL_18	SSTL_15
Data transfer frequency	The same as operational frequency	2 times that operational frequency	2 times that operational frequency	2 times that operational frequency

	DDR1 SDRAM	DDR2 SDRAM	DDR3 SDRAM
Data Rate	200 / 266 / 333 / 400 Mbps	400 / 533 / 677 / 800 Mbps	800 / 1066 / 1333 / 1600 Mbps
Supply Voltage	2.5V +/- 0.2 volts	1.8V +/- 0.1 volts	1.5V +/- 0.075 volts
Interface	SSTL_2	SSTL_18	SSTL_15
Data Strobe	Single ended	Single ended / Differential	Differential Default
Burst Length	2, 4, 8	4, 8	8
Prefetch	2	4	8
Number of Banks	4	4 / 8	8
Reset	No	No	Yes
On Die Termination	No	Yes	Yes, Dynamic ODT
Driver Calibration	No	Off-Chip Driver Calibration	Self Calibration w/ ZQ pin
Package	TSOP II	FBGA	FBGA

ODT : On-die termination (ODT) is the technology where the termination resistor for impedance matching in transmission lines is located inside a semiconductor chip instead of on a printed circuit board.

Types of Memory

Types of RAM in computer system

The following are some common types of RAM:

- **SRAM**: Static random access memory uses multiple transistors, typically four to six, for each memory cell but doesn't have a capacitor in each cell. It is used primarily for cache.
- **DRAM**: **Dynamic random access memory** has memory cells with a paired transistor and capacitor requiring constant refreshing.
- **FPM DRAM**: Fast page mode **dynamic random access memory** was the original form of DRAM. It waits through the entire process of locating a bit of data by column and row and then reading the bit before it starts on the next bit. Maximum transfer rate to L2 cache is approximately 176 MBps.
- **EDO DRAM**: **Extended data-out dynamic random access memory** does not wait for all of the processing of the first bit before continuing to the next one. As soon as the address of the first bit is located, EDO DRAM begins looking for the next bit. It is about five percent faster than FPM. Maximum transfer rate to L2 cache is approximately 264 MBps.
- **SDRAM**: **Synchronous dynamic random access memory** takes advantage of the burst mode concept to greatly improve performance. It does this by staying on the row containing the requested bit and moving rapidly through the columns, reading each bit as it goes. The idea is that most of the time the data needed by the CPU will be in sequence. SDRAM is about five percent faster than EDO RAM and is the most common form in desktops today. Maximum transfer rate to L2 cache is approximately 528 MBps.
- **DDR SDRAM**: **Double data rate synchronous dynamic RAM** is just like SDRAM except that is has higher bandwidth, meaning greater speed. Maximum transfer rate to L2 cache is approximately 1,064 MBps (for DDR SDRAM 133 MHZ).
- **RDRAM**: **Rambus dynamic random access memory** is a radical departure from the previous DRAM architecture. Designed by Rambus, RDRAM uses a **Rambus**

in-line memory module **(RIMM)**, which is similar in size and pin configuration to a standard DIMM. What makes RDRAM so different is its use of a special high-speed data bus called the Rambus channel. RDRAM memory chips work in parallel to achieve a data rate of 800 MHz, or 1,600 MBps. Since they operate at such high speeds, they generate much more heat than other types of chips. To help dissipate the excess heat Rambus chips are fitted with a heat spreader, which looks like a long thin wafer. Just like there are smaller versions of DIMMs, there are also SO-RIMMs, designed for notebook computers.

- Credit Card **Memory**: Credit card memory is a proprietary self-contained DRAM memory module that plugs into a special slot for use in notebook computers.
- **PCMCIA Memory Card**: Another self-contained DRAM module for notebooks, cards of this type are not proprietary and should work with any notebook computer whose system bus matches the memory card's configuration.
- **CMOS RAM**: CMOS RAM is a term for the small amount of memory used by your computer and some other devices to remember things like hard disk settings -- This memory uses a small battery to provide it with the power it needs to maintain the memory contents.
- **VRAM**: **VideoRAM**, also known as **multiport dynamic random access memory** (MPDRAM), is a type of RAM used specifically for video adapters or 3-D accelerators. The "multiport" part comes from the fact that VRAM normally has two independent access ports instead of one, allowing the CPU and graphics processor to access the RAM simultaneously. VRAM is located on the graphics card and comes in a variety of formats, many of which are proprietary. The amount of VRAM is a determining factor in the resolution and color depth of the display. VRAM is also used to hold graphics-specific information such as 3-D geometry data and texture maps. True multiport VRAM tends to be expensive, so today, many graphics cards use **SGRAM** (synchronous graphics RAM) instead. Performance is nearly the same, but SGRAM is cheaper.

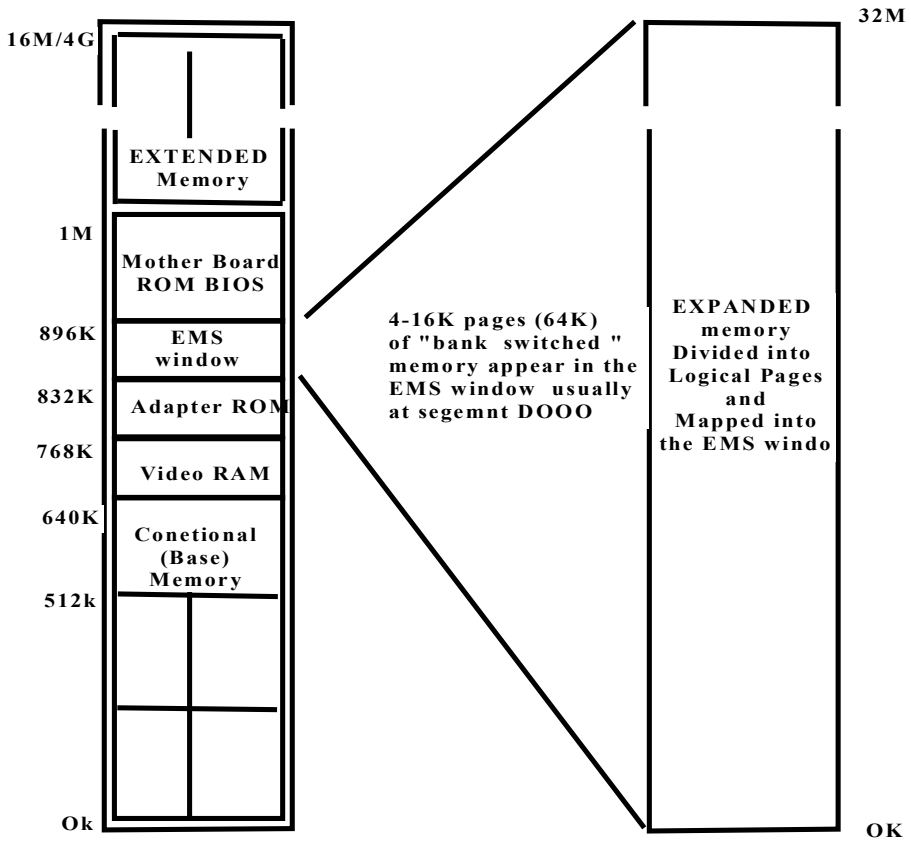
Logical Memory Organization , Conventional , Extended , Expanded Memory ,

Q.Explain the terms conventional , Extended and Expanded Memory

Ans.

The original PCs have total of 1MB of addressable memory and the top 384K of memory was reserved for the use by the system. Placing this reserved memory space at the top (between 640k and 1024k instead at the bottom between 0k and 640k) led to the what is called as ***Conventional memory barrier***. The different sections of memory in modern PCs are:

1) Conventional memory
2) Upper memory area
3) High memory area
4) Extended memory
5) Expanded memory
6) Video & RAM area
7) Adapter and special purpose ROM
8) Motherboard BIOS ROM.

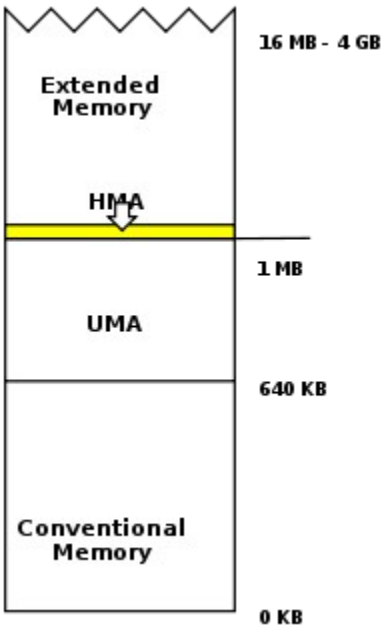


Conventional or Base memory: The original PC-XT type system was designed to use 1MB of memory space in RAM. This 1MB of RAM is divided into several sections, some of which have special use. Conventional memory is the memory area within first 1MB of system memory that can be used by DOS or application software. DOS can read write to the entire megabyte but can manage loading of program only in the portion of RAM space called Conventional Memory, which at the first PC was introduced was 512K. The other 512K was reserved for the use by the system itself including the mother board and adapter boards plugged into the system slots.

IBM decided after introducing the system in that only 384K was needed for the reserved uses , and company began marketing PCs with 640K of users memory. Thus 640K became the standard for memory that can be used by the DOS for running the program and is often called as ***memory barrier***.

First 640KB becomes the standard memory that can be used by DOS for running programs. This is also called as the base memory. After 640 K, some area is reserved for the use of graphics boards, other adapters and ROM BIOS of motherboard.

Extended memory : Extended memory is above 1MB. This memory is usually not available to computer directly. It can be accessed through special software or program or driver, this memory is available in AT machines, not in XTs. The limit of extended memory is 16MB for 286 processor and for 386, 486 and higher processor the extended memory is up to 4GB. System based on new Pentiums have a limit of 64G of memory.



The XMS (Extended Memory Specification was developed by Intel, Microsoft , Lotus Development and AST corp. in 1987 to specify how programs would use extended memory. This XMS Functions on system based on 286 and higher processor to use extended memory and another block of memory which out of the reach of the DOS. "HIMEM.SYS" is a driver to use extended memory This driver is loaded through the config.sys file . Extended memory is generally used by multitasking programs.

Extended memory can also be accessed directly by DOS programs running in protected mode using VCPI or DPMI, two (different and incompatible) methods of using protected mode under DOS.

Extended memory should not be confused with expanded memory, an earlier method for expanding the IBM PC's memory capacity beyond 640 kB (655,360 bytes) using an expansion card with bank switched memory modules. Because of the available support for expanded memory in popular applications, device drivers were developed that emulated expanded memory using extended memory. Later two additional methods were developed allowing direct access to a small portion of extended memory from real mode. These memory areas are referred to as the high memory area (HMA) and the upper memory area (UMA; also referred to as upper memory blocks or UMBs).

Expanded memory : Also known as *EMS (***E***xpanded* **M***emory* **S***pecification), expanded memory* is a technique for utilizing more than 1MB of main memory in DOS -based computers. The limit of 1MB is built into the DOS operating system. The upper 384K is reserved for special purposes, leaving just 640K of conventional memory for programs.

There are several versions of EMS. The original versions, called *EMS 3.0* and *3.2*, enable programs to use an additional 8MB of memory, but for data only. An improved version

developed by AST, Quadram and Ashton-Tate is known as *EEMS (Extended EMS)*. EEMS enables programs to use extra memory for code as well as for data. The most recent version of EMS (created in 1987) is known as *EMS 4.0* or *LIM 4.0*, LIM being the initials of the three companies that developed the specification: Lotus, Intel, and Microsoft. EMS 4.0 raises the available amount of memory to 32MB.

Expanded memory relies on technical trick known as paging. Paging involves accessing of large memory by swapping it in 16KB chunks or blocks in and out of 64KB memory window. This window is placed between unused memory space from 640KB to 1MB.Expanded memory is inefficient for the program code and is normally used for the data. This memory unlike conventional or extended this memory cannot be directly addressed by the processor.

Expanded memory is slow and clumsy for the system to use and generally obsolete (outdated) now days. But some of the antique software still require EMS memory. 386 and higher systems can create expanded memory out of extended memory out of extended memory by using a driver called "EMM386.EXE". This software was designed with 8-bit system because they do not have capability to access extended memory. This software is load through the Autoexec.bat file**.**

Until the release of Microsoft Windows 3.0 in 1990, expanded memory was the preferred way to add memory to a PC. The alternative method, called *extended memory*, was less flexible and could be used only by special programs such as RAM disks. Windows 3.0 and all later versions of Windows, however, contain an *extended memory manager* that enables programs to use extended memory without interfering with one another. In addition, Windows can simulate expanded memory for those programs that need it (by using the EMM386.EXE driver).

Main uses extended memory are:

RAM-disks

A RAM-disk is a chunk of semiconductor memory that behaves like an ordinary disk but is extremely fast. It also loses its data instantly once power is turned off but is great for temporary files such as index files, extracted data from Lotus to be imported into another application etc. Example: VDISK.SYS, supplied as part of DOS 3.x.

To create a RAM-disk, add one line in your config.sys such as DEVICE = \DOS\VDISK.SYS 256 /E. This will look in the DOS subdirectory for the driver, create a 256 kB RAM-disk. The /E parameter will place it into the extended memory.

Disk caches

A disk cache is a program to speed up disk access by storing the most frequently use information in the computer's memory and reading ahead from the disk in anticipation. With floppy disks, the time saved can be spectacular. Writes are almost always performed to the disk to prevent loss of data in case of power failure. Example: PC-CACHE, as supplied with PC-Tools. A shareware product is EMMCACHE.

Print spoolers

A print spooler utilises the computer's memory as a high speed buffer so that a fast computer is not slowed down by a slow printer. For example you can print a 100 page database report and then load a spread sheet program, print reports and graphs, then use your wordprocessor while the database report is still printing. Print spoolers that use extended memory usually come with the memory card. The AST SUPERSPL is a good example of a spooler with lots of options. A shareware product is EXTSPL.

Q.Difference between Extended and Expanded Memory

Ans. Extended memory

Memory addresses greater than or equal to one megabyte are called extended memory. The 8088 and 8086 PCs can't have extended memory because these chips can access only addresses of less than one megabyte (1MB) in size. With the minor exception of the High Memory Area (HMA), extended memory can be addressed only by applications run in real mode. It is possible, however, for DOS applications to make use of this memory to store data (but not to execute code directly from there). XMS (eXtended Memory Standard, promulgated by Microsoft) permits applications to allocate extended memory and takes care of copying data to and from extended memory and conventional memory so that the application does not have to worry about switching between modes.

Like EMS, XMS usually requires loading a device driver of some sort. Extended memory is limited to 15Mb on 286es and 386SXes (15Mb extended plus 1Mb conventional and upper memory equals 16Mb, or 224, 24 being the number of address lines coming out of the CPU), limited to 4 gigabytes (232) for 386DX chips and up, although very few motherboards have been designed to hold that much memory.

Expanded memory

Expanded memory is addressed from within the lower 1MB space, usually above 640K.

It is sometimes up to 64K of real addresses but this is just a small portion of the whole expanded memory, which can be very large. The expanded memory requires hardware and/or software that maps the expanded memory to a piece of address space, in what is called a "page frame". Extended memory can be used as expanded memory by using software and the 80286 or 80386 chips to "remap" it to the lower 1MB. It should be noted that the 80386 chip has hardware built in that supports expanded memory while the 80286 chip does not. Software that will convert extended memory on an 80286 machine to expanded memory may result in a significant performance penalty, if the machine does not have special hardware support for expanded memory. Software that will convert extended memory to expanded memory on an 80286 machine is not written to use the special hardware built into the 80386 chip, so the same type of performance penalty may apply.

Hardware supported expanded memory is the fastest form of expanded memory and is available directly on all 80386SX or better IBM compatibles running the proper software. It is also available on some 80286 machines with special chip sets or 80286 machines equipped with a hardware memory manager add-on. Lastly, it is available in a large number of memory expansion boards for all IBM compatible machines, including the 8088/8086 machines.

BIOS-Basic Input Output System Setup

Q.What is Bios? State it Functions

Ans. The BIOS (basic input output system) provides the processor with the information required to boot the system from a non volatile storage unit (HDD, FDD, Cdd or other). It provides the system with the settings and resources that are available on the system.

BIOS (Basic Input Output System) is an electronic set of instructions that a computer uses to successfully start operating. The BIOS is located on a chip inside of the computer and is designed in a way that protects it from disk failure.

The four main functions of a PC BIOS

- POST - Test the computer hardware and make sure no errors exist before loading the operating system. Additional information on the POST can be found on our POST and Beep Codes page.
- Bootstrap Loader - Locate the operating system. If a capable operating system is located, the BIOS will pass control to it.
- BIOS drivers - Low level drivers that give the computer basic operational control over your computer's hardware.

- BIOS or CMOS Setup - Configuration program that allows you to configure hardware settings including system settings such as computer passwords, time, and date.

The BIOS provides those **instructions**. Some of the other common tasks that the BIOS performs include:

- A power-on self-test (POST) for all of the different hardware components in the system to make sure everything is working properly
- Activating other BIOS chips on different cards installed in the computer - For example, SCSI and graphics cards often have their own BIOS chips.
- Providing a set of low-level routines that the operating system uses to interface to different hardware devices - It is these routines that give the BIOS its name. They manage things like the keyboard, the screen, and the serial and parallel ports, especially when the computer is booting.
- Managing a collection of settings for the hard disks, clock, etc.

The BIOS is special software that interfaces the major hardware components of your computer with the operating system. It is usually stored on a Flash memory chip on the motherboard, but sometimes the chip is another type of ROM.

When you turn on your computer, the BIOS does several things. This is its usual sequence:

1. Check the CMOS Setup for custom settings
2. Load the interrupt handlers and device drivers
3. Initialize registers and power management
4. Perform the power-on self-test (POST)
5. Display system settings
6. Determine which devices are bootable
7. Initiate the bootstrap sequence

Here is a list of top 5 most popular bios Manufacturers:

- AMI (American Megatrends, Inc.)
- Phoenix Technologies
- Dell
- Gateway
- IBM

Q.How It Is Stored on a Computer

- Traditional BIOS firmware is included on a computer motherboard as Read Only Memory (ROM). This means that the system is accessible but is not dynamically written to the way that a computer's hard drive is during operation. A battery on the motherboard keeps the data on it intact, even when the computer is turned off while not in use. Although sometimes confused with a Complementary Metal Oxide Semiconductor (CMOS), the BIOS refers to the firmware on the motherboard while the CMOS is the physical location where the date and system configuration data are stored.
- Many PC manufacturers today use flash-memory to hold this system, which allows users to more easily update it on computers. This can solve problems with the original BIOS or add new functionality to it. Users can periodically check for new versions, as some vendors release numerous updates over the course of a product's lifetime. To find an update, users should check the manufacturer of their motherboard. These updates should be done carefully, however, as incorrect versions or corrupted software can make a computer impossible to start up.

How Configuring and Controlling the BIOS is done

- PC users can make certain adjustments to the system through a configuration screen on the computer. The setup screen is typically accessed with a special key sequence during the first moments of startup, often "Delete" or a Function key. This setup screen allows users to change the order in which drives are accessed during startup, monitor computer component temperatures, and control the functionality of a number of devices. Most computer users never need to access these systems on their computer, though many fixes to simple errors can require adjustments.

Motherboard Selection Criteria

There are a wide variety of motherboards available today. When selecting a new mobo for your homebuilt computer, many things have to be taken into consideration, including:

- **Form Factor**. The form factor is a set of standards that include the size and shape of the board, the arrangement of the mounting holes, the power interface, and

the type and placement of ports and connectors. Generally, you should choose the case to fit the mobo, not vice-versa. But if there is a case that you simply *must*use (either because it's the one you happen to have or because you really, really like that case), then make sure the motherboard you choose is of a compatible form factor.

- **Processor support**. You must select a mobo that supports the type and speed of processor you want to use and has the correct type of socket for that processor.
- **RAM support**. Make sure that the motherboard you select supports enough RAM of the type (DDR-SDRAM, DDR2-SDRAM, RDRAM, etc.) that you want to use. Most motherboards manufactured as of this writing can support at least 4 Gig of RAM, with DDR2 being the most popular type because of its speed and relatively low cost. Most DDR motherboards also support *dual channel*DDR, which can further improve performance. But to take advantage of dual-channel, the RAM sticks must be installed in matched pairs, and the mobo must support it.
- **Chipset**. The chipset pretty much runs the show on the motherboard, and some chipsets are better than others. The chipset cannot be replaced, so the only way to solve problems caused by a bad chipset is to replace the mobo. Read the reviews of other motherboards using the same chipset as the one you are considering to see if a lot of people have reported problems with it.
- **SATA support**. There's really very little reason not to use SATA drives these days. They're priced comparably to EIDE drives, but deliver much higher data transfer. But to use SATA, your motherboard must have SATA support. (Well, you can actually install aftermarket SATA expansion cards, but why do that on a new computer?)
- **Expansion Slots and Ports**. How many of each type of expansion slot are included? Will they be enough to meet your current and future needs? How about Firewire support? And does it have enough USB slots for all the peripherals you want to dangle off of it?
- **Reputation**. Search the newsgroups to see if others have found the board you are considering to be a lemon. One excellent Web resource for motherboard research is Motherboards.org. When choosing a motherboard, reliability is the most important factor. Replacing a failed motherboard requires essentially disassembling the entire computer, and may also require reinstalling the operating system and applications from scratch.
- **Compatibility**. Most motherboards include drivers for all recent Windows versions, but check the documentation just to be sure. If you plan to use the board for a computer running another operating system (Linux, UNIX, BSD, etc.)

first check the with the motherboard manufacturer to see if it is compatible, and then search the hardware newsgroups for the OS you will be using to see how that particular board has worked out for others.

- **On-Board Features**. Do you want integrated audio or video? If you don't plan on using the computer for graphics, multimedia, or gaming, then you may be able to save money by buying a motherboard with less-than-spectacular integrated audio and/or video.
- **RAID Support**. RAID (Redundant Array of Independent Disks) is a set of protocols for arranging multiple hard drives into "arrays" to provide fault tolerance and/or increase the speed of data access from the hard drives. Many motherboards have RAID controllers built-in, saving you the cost of installing an add-on RAID controller.

Cost. Even if you are on a budget, the motherboard is not the place to cut corners. Try a less fancy case, instead. A good motherboard is more important than neon lights. But at the same time, the fact that one mobo costs twice as much as another doesn't mean it is twice as good. By searching newsgroups and reading hardware reviews, you're likely to find some inexpensive boards that perform as well as (or even better than) boards costing a great deal more.

Factor While selecting Mother With examples

Major Selection Criteria

While there may be a host of reasons why an individual selects a particular motherboard, there are only four major factors that one must consider. The following selection criteria should be used to narrow down the available motherboards so that the motherboard which is selected can be used for the intended application.

Processor

Before deciding on a motherboard, it is important to determine which type of CPU will be used. CPUs vary in size and pin configurations. Typically, a motherboard will work for a number of different CPUs, but not all will fit in the same motherboard. Intel and AMD, the two major CPU manufacturers, each use different pinouts for their chips, so a motherboard that works for one will not work for the other.

In addition, each of these manufactures uses several different socket pinouts. The socket must match the CPU exactly for the two to be able to work together. There is no sense in trying to provide a listing of which sockets fit which processors, as the available processors are constantly changing; however, the most common sockets are as follows:

Manufacturer	Socket	Used For
Intel	Socket 2011	Intel's most recent LGA socket. Used mostly by gamers for six core i7 processors.
Intel	Socket 1366	Workstation class computer using the i7 core or the Xenon 3XXX series. Has the pins on the motherboard.
Intel	Socket 1156	The average consumer socket for i3, i5, and i7 processors.
Intel	Socket 1155	A newer version of the socket 1156. Although it supports everything the 1156 does, it adds additional support for SATA III.
Intel	Socket 775	Intel's first LGA socket. Still a very popular socket. The first with the CPU pins on the motherboard.
Intel	Socket 771	The first Intel socket which allows for the use of dual processors. Used only for server applications.
AMD	Socket AM3	AMD's latest consumer socket. Same as AM2+ but uses only DDR3 memory.
AMD	Socket AM2+	The most common AMD socket. Supports both AM2 and AM3 processors.
AMD	Socket AM2	The oldest AMD socket in current production. Cannot support AM3 CPUs or DDR3 memory.
AMD	Socket F	The latest server socket by AMD. The first socket by AMD with the pins on the motherboard instead of on the CPU.

Motherboards are usually listed with the socket type as one of the specifications. CPUs are listed with the type of socket they require. Therefore, picking a motherboard which will work with a particular CPU mostly consists of checking on the sockets. When in doubt, it is best to check with the manufacturer of the motherboard in order to see which processors are compatible with it.

Form Factor

Not all motherboards are the same physical size. The most common size is ATX; however, the specification which created the ATX standard also provided for several alternate configurations. All ATX motherboards will have the same general configuration, with the major components located in the same places. While it is usually possible to mount a smaller motherboard into a computer case, mounting a larger one may be impossible.

A second motherboard configuration standard called "BTX" was created by Intel at a later time, in an attempt to solve airflow problems with the original ATX standard. However, this standard never caught on. While there are some BTX motherboards on the market, especially proprietary ones made by major computer manufacturers, they are not all that common.

The DTX standard was created by AMD as an answer to Intel's failed BTX standard.

Motherboard Form Factor	Size in Inches
Flex ATX	9.00 inches by 7.50 inches
Micro ATX/Embedded ATX	9.60 inches by 9.60 inches
Mini ATX	11.20 inches by 8.20 inches
Standard ATX	9.60 inches by 12.00 inches
Extended ATX (EATX)	12.00 inches by 13.00 inches
Workstation ATX (WATX)	14.00 inches by 16.75 inches
DTX	9.60 inches by 8.00 inches
Mini DTX	8.00 inches by 6.70 inches
ITX	8.46 inches by 7.50 inches
Mini ITX	6.70 inches by 6.70 inches
Nano ITX	4.70 inches by 4.70 inches
Pico ITX	3.90 inches by 2.80 inches

These dimensions are given only to help identify the type of motherboard currently in a computer. These motherboard standards not only affect the size of the motherboard, but also the configuration of the various connectors. Attempting to put an ITX motherboard into a computer case which previously housed an ATX motherboard can cause serious problems, as the card slots and connectors will not line up properly.

Memory Type

If one is replacing an existing motherboard and wants to reuse the existing memory modules, it is important to verify that the memory modules will fit into the new motherboard. Most modern computers use either DDR 2 orDDR 3 memory, with some of the older ones still using DDR (sometimes referred to as DDR1). Most will have four memory slots, configured in two banks of two slots. The other consideration for memory is speed. Not all motherboards and memory modules have the same speed.

Chipset

The chipset configures the motherboard and controls how the computer's CPU communicates with the rest of the computer. It also controls the bus speed of the motherboard; as such, it is vitally important. Chipsets will be rated by the speed at which they operate. Choosing a high-speed chipset allows faster memory to be used, and generally helps the computer run faster.

Please note that chipsets are not replaceable, but instead are a permanent part of the motherboard.

Besides features, the chipset also controls what features the motherboard has. Things like RAID control, surround sound, and support for USB 3 are all controlled by the chipset. The features of a motherboard will normally be listed, rather than specific information about the chipset.

Other Features and Options to Consider

While the four areas listed above are the major deciding factors when choosing any motherboard, especially to eliminate a motherboard from the running, there are a number of other features which can make the difference between one motherboard and another.

PCI Slots

The number and type of PCI extension card slots a motherboard has will affect the user's ability to add expansion cards. If the application for that computer requires a lot of expansion cards, then a motherboard needs to be selected which has room for them. As a general rule of thumb, the larger motherboard formats also have more expansion slots.

Input-Output (I/O) Connectors

Back in the early days of personal computers, most input and output functions were performed by expansion cards. Today's motherboards bring all of those functions on-board, eliminating the need for buying and installing additional cards. The number and types of connectors available can be critically important to some users and some applications. At a minimum, a motherboard will have the following:

- VGA monitor connection
- USB 2.0 ports
- 10/100 Ethernet connection
- Audio input and output connections

In addition, some motherboards may have any combination of the following:

- USB 3.0 port
- PS/2 mouse port
- PS/2 keyboard port
- Coaxial video jack
- Optical port
- Bluetooth transmitter
- eSATA ports
- Firewire port

While almost any computer peripheral will connect through a USB port, having these other connections can reduce the number of things connected through USB. Although as many as 127 devices can be connected to a single USB host controller, it must be remembered that the bandwidth of the bus must be shared between all those devices. Therefore, the more devices which can be directly connected through other ports, the better.

Onboard Networking

Almost all motherboards have onboard networking. This eliminates the need to add a card for the network connector. Some also have wireless network connectivity built-in.

Onboard Video

Almost all motherboards also have onboard video processing. This eliminates the need to add a separate video card. However, with heavy graphics applications or multiple

monitors, or to gain speed, a separate video card is still recommended. Graphics accelerator cards are especially useful for gaming.

SLI and Crossfire Compatibility

This is of special importance to high-end gamers or others who are using 3D graphics. SLI and Crossfire are the methods for using multiple video cards together to increase performance and quality. Most boards will have one or the other, although there are a few boards which are compatible with both.

Q.What is AGP state its Features?

Ans. The Accelerated Graphics Port (also called Advanced Graphics Port) is a high-speed point-to-point channel for attaching a graphics card to a computer's motherboard primarily to assist in the acceleration of 3D computer graphics.

AGP is often referred to as a bus'; however, this is a misnomer- a single AGP controller is only capable of controlling a single device.AGP is currently being phased out in favor of PCI Express. The ACP slot is designed to accept an AGP video card. This bus is in fact a port since it connects only two devices. Many motherboards with embedded video do not include an AGP slot. That means that upgrading the video adapter in such a configuration requires using a PCI video card.

Accelerated **G**raphics **P**ort) A high-speed 32-bit port from Intel for attaching a display adapter to a PC. It provides a direct connection between the card and memory, and only one AGP slot is on the motherboard. AGP was introduced as a higher-speed alternative to PCI display adapters, and it freed a PCI slot for another peripheral device. The brown AGP slot is slightly shorter than the white PCI slot and is located about an inch farther back. AGP was superseded by PCI Express.

Advantage of AGP over PCI

As computers increasingly became graphically oriented, successive generations of graphics adapters began to push the limits of PCI, a bus with shared bandwidth. This led to the development of AGP, a "bus" dedicated to graphics adapters.

The primary advantage of AGP over PCI is that it provides a dedicated pathway between the slot and the processor rather than sharing the PCI bus. In addition to a lack of contention for the bus, the direct connection allows for higher clock speeds. AGP also uses sideband addressing, meaning that the address and data buses are separated so the entire packet does not need to be read to get addressing information. This is

done by adding eight extra 8-bit buses which allow the graphics controller to issue new AGP requests and commands at the same time with other AGP data flowing via the main 32 address/data (AD) lines. This results in improved overall AGP data throughput.

In addition, to load a texture, a PCI graphics card must copy it from the system's RAM into the card's Video memory, whereas an AGP card is capable of reading textures directly from system RAM using the graphics address remapping table, which reapportions main memory as needed for texture storage, allowing the graphics card to access them directly. The maximum amount of system memory available to AGP is defined as the AGP aperture.

Data Transfer rate

- The AGP transfers axe 32 bits wide, but use 66.66 MHz clock speed.
- The AGP lx transferred I bit per data line per clock cycle yielding 266,66 Mbytes/s.
- The AGP 2x and 4x transferred 2 and 4 bits per data line per clock cycle yielding 533.33 and 1,06666 Mbytes/s respectively.
- The latest AGP 8 transfers 8 bits per data line per clock cycle yielding 2133.28 Mbytes/s. This implementation of AGP is
- already available in the latest motherboards and chipsets. AGP 8x is usually used in graphics intensive applications like video editing, 3-dimensional mapping, etc.

Feature of AGP

- ✓ **Peak Bandwidth** 4x the PCI bus, and higher sustained rates via sideband and pipelining.
- ✓ **Direct Memory Execute** of textures.
- ✓ **Reduced Contention** with the CPU and 1/0 devices for bus and memory access. The PCI bus serves disk controllers, LAN chips, and possibly video capture. AGP operates concurrently with and independent from, most transactions on PCI. Further, CPU accesses to system RAM can proceed concurrently with the graphics chip's AGP RAM reads, because of so-called out-of-order and queuing hardware support in the chip set, So inspite of bean- access from the graphics chip, there should be no audio breakup or other CPU degradation.
- ✓ **An "extra port"** to the graphics chip for memory access, so it can concurrently read textures from AGP memory while reading/writing Z-values and pixels from local memory.
- ✓ **Allowing the CPU** to write directly to shared system AGP memory when it needs to provide graphics data, such as commands or animated textures.

Generally the CPU can more quickly access main memory than it can graphics local memory via AGP, and certainly faster than via the PCI bus.

Q. What is the difference between an AGP and a PCI graphics card?

Answer: The biggest difference between AGP and PCI graphics cards is that AGP cards can access the system memory to help with complex operations such as texture mapping. PCI cards can only access the memory available on the actual card. AGP doesn't share bandwidth with other devices, whereas PCI cards do. AGP also makes pipelined requests, which means it can execute multiple instructions at one time. PCI cards are not pipelined, which means each instruction has to finish before the next one is run.

So, with all these great advantages of AGP, you'd think it would be the clear winner in performance, right? Well, not quite. Tests of similar AGP and PCI graphics cards show they perform almost the same (typically measured in frames per second). The area where AGP really shines is in high-resolution tests, where the direct access to the system memory is most beneficial.

If you're installing an AGP or PCI card in your computer, the AGP slot is usually the shortest and should be brown. The PCI slots are slightly longer and are colored white. The actual size of the cards can vary as much as a few inches, though the pins on the bottom of the card should match the correct slot.

Important: AGP and PCI slots are different sizes. Therefore, AGP cards can only be placed in AGP slots and PCI cards will only fit in PCI slots.

Q.Compare AGP and PCI

Ans. AGP will deliver a peak bandwidth that is 4 times higher than the PCI bus using pipelining, sideband addressing, and more data transfers per clock. It will also enable graphic cards to execute texture maps directly from system memory instead of forcing it to pre-load the texture data to the graphics card's local memory.

Features that set AGP apart from PCI

- Probably the most important feature of AGP is DIME (direct memory execute). This gives AGP chips the capability to access main memory directly for the complex operations of texture mapping.
- AGP provides the graphics card with two methods of directly accessing texture maps in system memory: pipelining and sideband addressing.
- AGP makes multiple requests for data during a bus or memory access, while PCI makes one request, and does not make another until the data it requested has been transferred.

- AGP doesn't share bandwidth with other devices, whereas the PCI bus does share bandwidth.

	AGP	**PCI**
1	Pipelined Requests	Non-pipelined
2	Address/Data de-multiplexed	Address/Data multiplexed
3	Peak at 533 MB/s in32 bits	Peak at 133 MB/s in 32 bits
4	Single target, single master	Multi target, multi master
5	Memory read/write only No other input/output operations	Link to entire system
6	High/low priority queues	No priority queues

Q. What does all this mean?

DIME	short for Direct Memory Execute, DIME allows for video card to use some of the main memory for texture memory with 3D graphics. Usually video cards have 4 MB of RAM, some have 8 MB of RAM, but DIME allows for 12, 16, or even more memory to be used by allocating some of the main system memory.
Pipelining	As you should know from reading Hennessy and Patterson's 'great' Computer Architecture book, pipelining is an implementation technique whereby multiple instructions are overlapped in execution. A pipeline is just like an assembly line. There are various different steps (pipe stage or pipe segments) that contribute to the end result. Each of these steps are done in parallel. The opposite of a pipelined architecture is a sequential architecture, in which steps are completed sequentially or one after another, not in parallel.
Sideband addressing	the AGP bus uses sideband signals to send addressing information separately from data. This technique allows addresss informaton to be presented to the bus concurrent with a data transaction. The result is a more efficient use of the AGP bus for data transfers. With sideband addressing, AGP utilized 8 extra "sideband lines" which allow the graphics controller to issue new addresses and requests simulataneously while data continues to move from previous requests on the main 32 data/address wires.
Bandwidth	the amount of data a network can transport in a certain period of time - it is the capacity for the rate of transfer, which is usually expressed in bits per second.

- AGP's is pipelined therefore requests are executed in parallel, making execution faster than the non-pipelined PCI bus.
- AGP's address/data is de-muxed, therefore the AGP pipeline can work with the data gotten from the de-mux. The PCI bus's address/data remained muxed so that the non-pipelined PCI bus works with the data gotten from the mux.

Further comparisons between AGP and PCI

- AGP is a port (it only connects two nodes) while PCI is a bus
- AGP does not replace the PCI bus, it is a dedicated connection that can be used only by the graphics subsystem
- AGP and PCI also differ in terms of their minimum length and alignment requirements for transactions. AGP transaction are multiples of 8 bytes in length and are aligned on 8 byte boundaries, while PCI transactions must be multiples of 4 bytes and are aligned on 4 byte boundaries.

Latest Chipset for Motherboard

CMOS Setup

Additional - Information

Q.State Feature of XT mother Board

Ans. The following list shows the highlights of these new models:

- Enhanced keyboard standard on Models 268, 278, and 089; 101 keys, and no status LEDs
- (XT interface cannot drive LEDs)
- Standard PC XT keyboard on Models 267, 277, and 088
- More disk capacity (20MB)
- Standard 5 1/4-inch, half-height, 360KB floppy drive
- Available 3 1/2-inch, half-height, 720KB floppy drive
- Capacity for four half-height storage devices within the system unit
- Capacity to expand to 640KB memory on system board without using expansion slots

System Architecture

- Microprocessor 8088
- Clock speed 4.77MHz
- Bus type ISA (Industry Standard Architecture)
- Bus width 8-bit
- Interrupt levels 8 (6 usable)
- Type Edge-triggered
- Shareable No
- DMA channels 4 (3 usable)
- Bus masters supported No
- Upgradeable processor complex No

Memory Architechture

- Standard on system board 256KB or 640KB
- Maximum on system board 256KB or 640KB
- Maximum total memory 640KB
- Memory speed (ns) and type 200ns Dynamic RAM chips
- System board memory-socket type 16-pin DIP
- Number of memory-module sockets 36 (4 banks of 9)
- Memory used on system board 36 64KB ×1-bit DRAM chips in 4 banks of 9, or 2 banks of 9
- 256KB ×1-bit and 2 banks of 9 64KB ×1-bit chips
- Memory cache controller No

- Wait states:
- System board 1
- Adapter 1

Standard Feature

- ROM size 40KB or 64KB
- ROM shadowing No
- Optional math coprocessor 8087
- Coprocessor speed 4.77MHz
- Standard graphics None standard
- RS232C serial ports 1 (some models)
- UART chip used NS8250B
- Maximum speed (bits per second) 9,600bps
- Maximum number of ports supported 2
- Pointing device (mouse) ports None standard
- Parallel printer ports 1 (some models)
- Bidirectional No
- Maximum number of ports supported 3
- CMOS real-time clock (RTC) No
- CMOS RAM None

Expansion Slot

- Total adapter slots 8
- Number of long/short slots 6/2
- Number of 8-/16-/32-bit slots 8/0/0
- Available slots (with video) 4

Q.State Feature of AT mother Board

Ans. Every system unit for AT models has these major functional components:

- ■ Intel 80286 (6MHz or 8MHz) microprocessor
- ■ Socket for 80287 math coprocessor
- ■ Eight I/O expansion slots (six 16-bit, two 8-bit)
- ■ 256KB of Dynamic RAM (base model)
- ■ 512KB of Dynamic RAM (enhanced models)
- ■ ROM-based diagnostics (POST)
- ■ BASIC language interpreter in ROM
- ■ Hard/floppy disk controller
- ■ 1.2MB hard disk floppy drive

- 20MB or 30MB hard disk drive (enhanced models)
- Serial/parallel interface (enhanced models)
- CMOS clock-calendar and configuration with battery backup
- Keylock
- 84-key keyboard or 101-key Enhanced keyboard (standard on newer models)
- Switchable worldwide power supply

System Architecture

- Microprocessor 80286
- Clock speed 6MHz or 8MHz
- Bus type ISA (Industry Standard Architecture)
- Bus width 16-bit
- Interrupt levels 16 (11 usable)
- Type Edge-triggered
- Shareable No
- DMA channels 8 (7 usable)
- Bus masters supported Yes
- Upgradeable processor complex No

Memory Architecture

- Standard on system board 512KB
- Maximum on system board 512KB
- Maximum total memory 16MB
- Memory speed (ns) and type 150ns Dynamic RAM chips
- System board memory-socket type 16-pin DIP
- Number of memory-module sockets 18 or 36 (2 or 4 banks of 18)
- Memory used on system board 36 128KB ×1-bit DRAM chips in 2 banks of 18, or
- 18 256KB ×1-bit chips in one bank
- Memory cache controller No
- Wait states:
- System board 1
- Adapter 1

Standard Features

- ROM size 64KB
- ROM shadowing No
- Optional math coprocessor 80287

- Coprocessor speed 4 or 5.33MHz
- Standard graphics None standard
- RS232C serial ports 1 (some models)
- UART chip used NS16450
- Maximum speed (bits per second) 9,600bps
- Maximum number of ports supported 2
- Pointing device (mouse) ports None standard
- Parallel printer ports 1 (some models)
- Bidirectional Yes
- Maximum number of ports supported 3
- CMOS real-time clock (RTC) Yes
- CMOS RAM 64 bytes
- Battery life 5 years

Expansion Slots

- Total adapter slots 8
- Number of long and short slots 8/0
- Number of 8-/16-/32-bit slots 2/6/0
- Available slots (with video) 5

Q.Differences Between PC/XT and AT Systems

Ans.

Systems that feature an 8-bit memory bus are called PC/XT systems after the pioneering IBM PC and IBM PC/XT. As you can see in Table 2.2, the differences between these systems and descendents of the IBM AT (16-bit memory bus and above) are significant. All modern systems fall into the AT category.

Table . Differences Between PC/XT and AT Systems

System Attributes PC/XT Type	8-Bit	16-, 32-, 64-Bit AT Type
Supported processors	All x86 or x88	286 or higher
Processor modes	Real	Real, Protected, Virtual Real[2]
Software supported	16-bit only	16- or 32-bit[2]
Bus slot width	8-bit	16-, 32-[1], and 64-bit[4]
Slot type	ISA only	ISA, EISA[1], MCA, PC Card, Cardbus[3], VL-Bus[3], PCI[3], AGP[4]
Hardware interrupts	8 (6 usable)	16 (11 usable)

Table . Differences Between PC/XT and AT Systems		
System Attributes	PC/XT Type	16-, 32-, 64-Bit AT Type
DMA channels	4 (3 usable)	8 (7 usable)
Maximum RAM	1MB	16MB/4GB[1] or more
Floppy controller speed	250Kbps	250, 300, 500, and 1000Kbps
Standard boot drive	360KB or 720KB	1.2M, 1.44MB, and 2.88MB
Keyboard interface	Unidirectional	Bidirectional
CMOS memory/clock	None standard	MC146818-compatible
Serial-port UART	8250B	16450/16550A or better

Storage Devices & Interfacing.

Objective

- To understand the Recording techniques in storage devices.
- To understand the working of storage devices.

2.1 Recording Techniques: FM, MFM , RLL, perpendicular recording

2.2 Hard Disk construction and working.

2.3 Terms related to Hard Disk.

Track, sector, cylinder, cluster, landing zone, MBR, zone recording, write pre-compensation.

2.4 Formatting: Low level, High level & partitioning.

2.5 FAT Basics: Introduction to file system, FAT 16, FAT 32, NTFS,

2.6 Hard Disk Interface: Features of IDE, SCSI, PATA, SATA, Cables & Jumpers.

2.7 CD ROM Drive: Construction, recording.(Block diagram)

2.8 DVD: Construction, Recording. (Block Diagram)

2.9 Blue-ray Disc specification.

Display Devices & Interfacing

Objective

 To understand the construction and working of display devices like CRT, LCD.

 To understand the Interfacing of above devices to PC.

3.1 CRT: - Block diagram & working of monochrome & colour Monitor

3.2 Characteristics of CRT Monitor :-

DOT Pitch, Resolution, Horizontal Scanning frequency, Vertical scanning frequency, Interlaced Scanning, Non-Interfaced scanning, Aspect ratio.

3.3 LCD Monitor: - Functional Block Diagram of LCD monitor, working principle, Passive matrix, Active matrix LCD display.

3.4 Touch Screen Display – The construction and working principle

3.4 Plasma Display Technology: - Construction & working principle.

3.5 Basic Block Diagram of Video Accelerator card